

Asymptotic Inference of Autocovariances of Stationary Processes

Han Xiao and Wei Biao Wu

Department of Statistics

5734 S. University Ave.

Chicago, IL 60637

e-mail: xiao@galton.uchicago.edu

e-mail: wbwu@galton.uchicago.edu

Abstract: The paper presents a systematic theory for asymptotic inference of autocovariances of stationary processes. We consider nonparametric tests for serial correlations based on the maximum (or \mathcal{L}^∞) and the quadratic (or \mathcal{L}^2) deviations. For these two cases, with proper centering and rescaling, the asymptotic distributions of the deviations are Gumbel and Gaussian, respectively. To establish such an asymptotic theory, as byproducts, we develop a normal comparison principle and propose a sufficient condition for summability of joint cumulants of stationary processes. We adopt a simulation-based block of blocks bootstrapping procedure that improves the finite-sample performance.

AMS 2000 subject classifications: Primary 60F05, 62M10; secondary 62E20.

Keywords and phrases: Autocovariance, blocks of blocks bootstrapping, Box-Pierce test, extreme value distribution, moderate deviation, normal comparison, physical dependence measure, short range dependence, stationary process, summability of cumulants.

1. Introduction

If $(X_i)_{i \in \mathbb{Z}}$ is a real-valued stationary process, then from a second-order inference point of view it is characterized by its mean $\mu = \mathbb{E}X_i$ and the autocovariance function $\gamma_k = \mathbb{E}[(X_0 - \mu)(X_k - \mu)]$, $k \in \mathbb{Z}$. Assume $\mu = 0$. Given observations X_1, \dots, X_n , the natural estimates of γ_k and the autocorrelation $r_k = \gamma_k/\gamma_0$ are

$$\hat{\gamma}_k = (1/n) \sum_{i=|k|+1}^n X_{i-|k|} X_i \quad \text{and} \quad \hat{r}_k = \hat{\gamma}_k / \hat{\gamma}_0, \quad 1 - n \leq k \leq n - 1, \quad (1)$$

respectively. The estimator $\hat{\gamma}_k$ plays a crucial role in almost every aspect of time series analysis. It is well-known that for linear processes with *independent and identically distributed* (iid) innovations, under suitable conditions, $\sqrt{n}(\hat{\gamma}_k - \gamma_k) \Rightarrow \mathcal{N}(0, \tau_k^2)$, where \Rightarrow stands for convergence in distribution, $\mathcal{N}(0, \tau_k^2)$ denotes the normal distribution with mean zero and variance τ_k^2 . Here τ_k^2 can be calculated by Bartlett's formula (see Section 7.2 of Brockwell and Davis (1991)). Other contributions on linear processes include Hannan and Heyde (1972), Hosoya and Taniguchi (1982), Anderson (1991) and Phillips and Solo (1992) etc. Romano and Thombs (1996) and Wu (2009) considered the asymptotic normality of $\hat{\gamma}_k$ for nonlinear processes. As a

primary goal of the paper, we shall study asymptotic properties of the quadratic (or \mathcal{L}^2) and the maximum (or \mathcal{L}^∞) deviations of $\hat{\gamma}_k$.

1.1. The \mathcal{L}^2 Theory

Testing for serial correlation has been extensively studied in both statistics and econometrics, and it is a standard diagnostic procedure after a model is fitted to a time series. Classical procedures include Durbin and Watson (1950, 1951), Box and Pierce (1970), Robinson (1991) and their variants. The Box-Pierce portmanteau test uses $Q_K = n \sum_{k=1}^K \hat{r}_k^2$ as the test statistic, and rejects if it lies in the upper tail of χ_K^2 distribution. An arguable deficiency of this test and many of its modified versions (for a review see for example Escanciano and Lobato (2009)) is that the number of lags K included in the test is held as a constant in the asymptotic theory. As commented by Robinson (1991):

"...unless the statistics take account of sample autocorrelations at long lags there is always the possibility that relevant information is being neglected..."

The problem is particularly relevant if practitioners have no prior information about the alternatives. The attempt of incorporating more lags emerged naturally in the spectral domain analysis; see among others Durlauf (1991), Hong (1996) and Deo (2000). The normalized spectral density $f(\omega) = (2\pi)^{-1} \sum_{k \in \mathbb{Z}} r_k \cos(k\omega)$ should equal to $(2\pi)^{-1}$ when the serial correlation is not present. Let $\hat{f}(\omega) = \sum_{k=1-n}^{n-1} h(k/s_n) \hat{r}_k \cos(k\omega)$ be the lag-window estimate of the normalized spectral density, where $h(\cdot)$ is a kernel function and s_n is the bandwidth satisfying the natural condition $s_n \rightarrow \infty$ and $s_n/n \rightarrow 0$. The former aims to include correlations at large lags. A test for the serial correlation can be obtained by comparing \hat{f} and the constant function $f(\omega) \equiv (2\pi)^{-1}$ using a suitable metric. In particular, using the quadratic metric and rectangle kernel, the resulting test statistic is the Box-Pierce statistic with unbounded lags. Hong (1996) established the following result:

$$\frac{1}{\sqrt{2s_n}} \left(n \sum_{k=1}^{s_n} (\hat{r}_k - r_k)^2 - s_n \right) \Rightarrow \mathcal{N}(0, 1), \quad (2)$$

under the condition that X_i are iid, which implies that all r_k in the preceding equation are zero. Lee and Hong (2001) and Duchesne, Li and Vandermeersch (2010) studied similar tests in spectral domain, but using a wavelet basis instead of trigonometric polynomials in estimating the spectral density and henceforth working on wavelet coefficients. Fan (1996) considered a similar problem in a different context and proposed *adaptive Neyman* test and thresholding tests, using $\max_{1 \leq k \leq s_n} (Q_k - k)/\sqrt{2k}$ and $n \sum_{k=1}^{s_n} \hat{r}_k^2 I(|\hat{r}_k| > \delta)$ as test statistics respectively, where δ is a threshold value. Escanciano and Lobato (2009) proposed to use Q_{s_n} with s_n being selected by AIC or BIC.

It has been an important and difficult question on whether the iid assumption in Hong (1996) can be relaxed. Similar problems have been studied by Durlauf (1991), Deo (2000) and Hong and Lee (2003) for the case that X_i are martingale differences. Recently Shao (2011) showed that (2) is true when (X_i) is a general white noise sequence, under the geometric moment contraction (GMC) condition. Since the GMC condition, which implies that the autocovariances decay geometrically, is quite strong, the question arises

as to whether it can be replaced by a weaker one. Furthermore, one may naturally ask: what if the serial correlation is present in (2)? To the best of our knowledge, there has been no results in the literature for this problem. This paper shall address these questions and substantially generalizes earlier results. We shall prove that (2) remains true even if all or some of r_k are not zero, but the variance of the limiting distribution, being different, will depend on the values of r_k . Furthermore, we derive the limiting distribution of $\sum_{k=1}^{s_n} \hat{r}_k^2$ when the serial correlation is present. The latter result enables us to calculate the asymptotic power of the Box-Pierce test with unbounded lags.

1.2. The \mathcal{L}^∞ Theory

Another natural omnibus choice is to use the maximum autocorrelation as the test statistic. Wu (2009) obtained a stochastic upper bound for

$$\sqrt{n} \max_{1 \leq k \leq s_n} |\hat{\gamma}_k - \gamma_k|, \quad (3)$$

and argued that in certain situations the test based on (3) has a higher power over the Box-Pierce tests with unbounded lags in detecting weak serial correlation. It turns out that the uniform convergence of autocovariances is also closely related to the estimation of orders of ARMA processes or linear systems in general. The pioneer works in this direction were given by E. J. Hannan and his collaborators, see for example Hannan (1974) and An, Chen and Hannan (1982). For a summary of these works we recommend (Hannan and Deistler, 1988, Section §5.3) and references therein. In particular, An, Chen and Hannan (1982) showed that if $s_n = O[(\log n)^\alpha]$ for some $\alpha < \infty$, then with probability one

$$\sqrt{n} \max_{1 \leq k \leq s_n} |\hat{\gamma}_k - \gamma_k| = O(\log \log n). \quad (4)$$

The question of deriving the asymptotic distribution of (3) is more challenging. Although Wu (2009) was not able to obtain the limiting distribution of (3), his work provided important insights into this problem. Assuming $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and $h \geq 0$, he showed that, for $T_k = \sqrt{n}(\hat{\gamma}_k - \mathbb{E}\hat{\gamma}_k)$,

$$(T_{k_n}, T_{k_n+h})^\top \Rightarrow \mathcal{N} \left[0, \begin{pmatrix} \sigma_0 & \sigma_h \\ \sigma_h & \sigma_0 \end{pmatrix} \right], \quad \text{where } \sigma_h = \sum_{k \in \mathbb{Z}} \gamma_k \gamma_{k+h}, \quad (5)$$

and we use the superscript \top to denote the transpose of a vector or a matrix. The asymptotic distribution in (5) does not depend on the speed of $k_n \rightarrow \infty$. It suggests that, at large lags, the covariance structure of (T_k) is asymptotically equivalent to that of the Gaussian sequence

$$(G_k) := \left(\sum_{i \in \mathbb{Z}} \gamma_i \eta_{i-k} \right) \quad (6)$$

where η_i 's are iid standard normal random variables. Define the sequences (a_n) and (b_n) as

$$a_n = (2 \log n)^{-1/2} \quad \text{and} \quad b_n = (2 \log n)^{1/2} - (8 \log n)^{-1/2} (\log \log n + \log 4\pi). \quad (7)$$

According to Berman (1964) (also see Remarks 3 and 4), under the condition $\lim_{n \rightarrow \infty} \mathbb{E}(G_0 G_n) \log n = 0$,

$$\lim_{s \rightarrow \infty} P \left(\max_{1 \leq i \leq s} |G_i| \leq \sqrt{\sigma_0} (a_{2s} x + b_{2s}) \right) = \exp\{-\exp(-x)\}.$$

Therefore, Wu (2009) conjectured that under suitable conditions, one has the Gumbel convergence

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq k \leq s_n} |T_k| \leq \sqrt{\sigma_0} (a_{2s_n} x + b_{2s_n}) \right) = \exp\{-\exp(-x)\}. \quad (8)$$

In a recent work, Jirak (2011) proved this conjecture for linear processes and for s_n growing with at most logarithmic speed. We shall prove (8) in Section 4 for general stationary processes; and our result allows s_n to grow as $s_n = O(n^\eta)$ for some $0 < \eta < 1$, and η can be arbitrarily close to 1 under appropriate moment and dependence conditions. The latter result substantially relaxes the severe restriction on the growth speed in (4) and Jirak (2011) and, moreover, the obtained distributional convergence are more useful for statistical inference. For example, other than testing for serial correlation and estimating the order of a linear system, (8) can also be used to construct simultaneous confidence intervals of autocovariances.

1.3. Relations with the Random Matrix Theory

In a companion paper, using the asymptotic theory of sample autocovariances developed in this paper, Xiao and Wu (2010) studied convergence properties of estimated covariance matrices which are obtained by banding or thresholding. Their bounds are analogs under the time series context to those of Bickel and Levina (2008a,b). There is an important difference between these two settings: we assume that only one realization is available, while Bickel and Levina (2008a,b) require multiple iid copies of the underlying random vector.

There has been some related works in the random matrix theory literature that are similar to (8). Suppose one has n iid copies of a p -dimensional random vector, forming a $p \times n$ data matrix \mathbf{X} . Let \hat{r}_{ij} , $1 \leq i, j \leq p$, be the sample correlations. Jiang (2004) showed that the limiting distribution of $\max_{1 \leq i < j \leq p} |\hat{r}_{ij}|$, after suitable normalization, is Gumbel provided that each column of \mathbf{X} consists of p iid entries and each entry has finite moment of some order higher than 30, and p/n converges to some constant. His work was followed and improved by Zhou (2007) and Liu, Lin and Shao (2008). In a recent article, Cai and Jiang (2010) extended those results in two ways: (i) the dimension p could grow exponentially as the sample size n provided exponential moment conditions; and (ii) they showed that the test statistic $\max_{|i-j| > s_n} |\hat{r}_{ij}|$ also converges to the Gumbel distribution if each column of \mathbf{X} is Gaussian and is s_n -dependent. The latter generalization is important since it is one of the very few results that allow dependent entries. Their method is Poisson approximation (see for example Arratia, Goldstein and Gordon, 1989), which heavily depends on the fact that for each sample correlation to be considered, the corresponding entries are independent. Schott (2005) proved that $\sum_{1 \leq i < j \leq p} \hat{r}_{ij}^2$ converges to normal distribution after suitable normalization, under the conditions that each column of \mathbf{X} contains iid Gaussian entries and p/n converges to some positive constant. His proof heavily depends on the normality assumption. Techniques developed in those papers are not applicable here since we have *only one realization* and the dependence structure among the entries can be quite complicated.

1.4. A Summary of Results of the Paper

We present the main results in Section 2, which include a central limit theory of (2) and the Gumbel convergence (8). The proofs are given in Section 4. In Section 5 we prove a normal comparison principle, which is of independent interest. Since summability conditions of joint cumulants are commonly used in time series analysis (see for example Brillinger (2001) and Rosenblatt (1985)) and is needed in the proof of Theorem 4, we present a sufficient condition in Section 6. Some auxiliary lemmas are collected in Section 7. We also conduct a simulation study in Section 3, where we design a simulation-based block of blocks bootstrapping procedure that improves the finite-sample performance.

2. Main Results

To develop an asymptotic theory for time series, it is necessary to impose suitable measures of dependence and structural assumptions for the underlying process (X_i) . Here we shall adopt the framework of Wu (2005). Assume that (X_i) is a stationary causal process of the form

$$X_i = g(\cdots, \epsilon_{i-1}, \epsilon_i), \quad (9)$$

where $\epsilon_i, i \in \mathbb{Z}$, are iid random variables, and g is a measurable function for which X_i is a properly defined random variable. For notational simplicity we define the operator Ω_k : suppose $X = h(\epsilon_j, \epsilon_{i-1}, \dots)$ is a random variable which is a function of the innovations $\epsilon_l, l \leq j$, then $\Omega_k(X) := h(\epsilon_j, \dots, \epsilon_{k+1}, \epsilon'_k, \epsilon_{k-1}, \dots)$, where $(\epsilon'_k)_{k \in \mathbb{Z}}$ is an iid copy of $(\epsilon_k)_{k \in \mathbb{Z}}$. Namely ϵ_k in X is replaced by ϵ'_k .

For a random variable X and $p > 0$, we write $X \in \mathcal{L}^p$ if $\|X\|_p := (\mathbb{E}|X|^p)^{1/p} < \infty$, and in particular, use $\|X\|$ for the \mathcal{L}^2 -norm $\|X\|_2$. Assume $X_i \in \mathcal{L}^p, p > 1$. Define the *physical dependence measure of order p* as

$$\delta_p(i) = \|X_i - \Omega_0(X_i)\|_p, \quad (10)$$

which quantifies the dependence of X_i on the innovation ϵ_0 . Our main results depend on the decay rate of $\delta_p(i)$ as $i \rightarrow \infty$. Let $p' = \min(2, p)$ and define

$$\begin{aligned} \Theta_p(n) &= \sum_{i=n}^{\infty} \delta_p(i), \quad \Psi_p(n) = \left(\sum_{i=n}^{\infty} \delta_p(i)^{p'} \right)^{1/p'}, \quad \text{and} \\ \Delta_p(n) &= \sum_{i=0}^{\infty} \min\{\mathcal{C}_p \Psi_p(n), \delta_p(i)\}, \end{aligned} \quad (11)$$

where \mathcal{C}_p is defined in (30). It is easily seen that $\Psi_p(\cdot) \leq \Theta_p(\cdot) \leq \Delta_p(\cdot)$. We use Θ_p, Ψ_p and Δ_p as shorthands for $\Theta_p(0), \Psi_p(0)$ and $\Delta_p(0)$ respectively. We make the convention that $\delta_p(k) = 0$ for $k < 0$.

There are several reasons that we use the framework (9) and the dependence measure (10). First, the class of processes that (9) represents is huge and it includes linear processes, bilinear processes, Volterra processes, and many other time series models. See, for instance, Tong (1990) and Wiener (1958). Second, the physical dependence measure is easy to work with and it is directly related to the underlying data-generating mechanism. Third, it enables us to develop an asymptotic theory for complicated statistics of time series.

2.1. Maximum deviations of sample autocovariances

Note that $\hat{\gamma}_k$ is a biased estimate of γ_k with $\mathbb{E}\hat{\gamma}_k = (1 - |k|/n)\gamma_k$. It is then more convenient to consider the centered version $\max_{1 \leq k \leq s_n} \sqrt{n}|\hat{\gamma}_k - \mathbb{E}\hat{\gamma}_k|$ instead of $\max_{1 \leq k \leq s_n} \sqrt{n}|\hat{\gamma}_k - \gamma_k|$. Recall (7) for a_n and b_n .

Theorem 1. Assume $\mathbb{E}X_i = 0$, $X_i \in \mathcal{L}^p$ for some $p > 4$, and $\Theta_p(m) = O(m^{-\alpha})$, $\Delta_p(m) = O(m^{-\alpha'})$ for some $\alpha \geq \alpha' > 0$. If s_n satisfies $s_n \rightarrow \infty$ and $s_n = O(n^\eta)$ with

$$0 < \eta < 1, \quad \eta < \alpha p/2, \quad \text{and} \quad \eta \min\{2(p-2-\alpha p), (1-2\alpha')p\} < p-4, \quad (12)$$

then for all $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq k \leq s_n} |\sqrt{n}[\hat{\gamma}_k - (1 - k/n)\gamma_k]| \leq \sqrt{\sigma_0}(a_{2s_n}x + b_{2s_n}) \right) = \exp\{-\exp(-x)\}. \quad (13)$$

In (12), if $p \leq 2 + \alpha p$ or $1 \leq 2\alpha'$, then the second and third conditions are automatically satisfied, and hence Theorem 1 allows a very wide range of lags $s_n = O(n^\eta)$ with $0 < \eta < 1$. In this sense Theorem 1 is nearly optimal.

For the maximum deviation $\max_{1 \leq k < n} |\hat{\gamma}_k - \mathbb{E}\hat{\gamma}_k|$ over the whole range $1 \leq k < n$, it seems not possible to derive a limiting distribution by using our method. However, we can obtain a sharp bound $(n^{-1} \log n)^{1/2}$. The upper bound is given in (15), while the lower bound can be obtained by applying Theorem 1 and choosing a sufficiently small η such that (12) holds. Using Theorem 2, Xiao and Wu (2010) derived convergence rates for the thresholded autocovariance matrix estimates.

Theorem 2. Assume $\mathbb{E}X_i = 0$, $X_i \in \mathcal{L}^p$ for some $p > 4$, and $\Theta_p(m) = O(m^{-\alpha})$, $\Delta_p(m) = O(m^{-\alpha'})$ for some $\alpha \geq \alpha' > 0$. If

$$\alpha > 1/2 \quad \text{or} \quad \alpha' p > 2 \quad (14)$$

then for $c_p = 6(p+4)e^{p/4}\kappa_4\Theta_4$,

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq k < n} |\hat{\gamma}_k - \mathbb{E}\hat{\gamma}_k| \leq c_p \sqrt{\frac{\log n}{n}} \right) = 1. \quad (15)$$

Since $\Theta_p(m) \geq \Psi_p(m)$, we can assume $\alpha \geq \alpha'$. For a detailed discussion on their relationship, see Remark 6 of Xiao and Wu (2010). It turns out that for the special case of linear processes the condition (12) can be weakened to the following one:

$$0 < \eta < 1, \quad \eta < \alpha p/2, \quad \text{and} \quad (1-2\alpha)\eta < (p-4)/p. \quad (16)$$

See Remark 2. Furthermore, for linear processes the condition (14) can be relaxed to $\alpha p > 2$ as well.

In practice, the mean $\mu = \mathbb{E}X_0$ is often unknown and we can estimate it by the sample mean $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$. The usual estimates of autocovariances and autocorrelations are

$$\check{\gamma}_k = \frac{1}{n} \sum_{i=k+1}^n (X_{i-k} - \bar{X}_n)(X_i - \bar{X}_n) \quad \text{and} \quad \check{\gamma}_k = \check{\gamma}_k / \check{\gamma}_0. \quad (17)$$

Corollary 3. *Theorem 1 and Theorem 2 still hold if we replace $\hat{\gamma}_k$ therein by $\check{\gamma}_k$. Furthermore,*

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq k \leq s_n} |\sqrt{n}[\check{\gamma}_k - (1 - k/n)r_k]| \leq (\sqrt{\sigma_0}/\gamma_0)(a_{2s_n}x + b_{2s_n}) \right) = \exp\{-\exp(-x)\}.$$

Proof of Corollary 3. For the $\check{\gamma}_k$ version of Theorem 1, it suffices to show that

$$\max_{1 \leq k \leq s_n} |\sqrt{n}(\check{\gamma}_k - \hat{\gamma}_k)| = o_P \left(\frac{1}{\sqrt{\log s_n}} \right). \quad (18)$$

Let $S_k = \sum_{i=1}^k X_i$. By Theorem 1 (iii) of Wu (2007), we have $\|\max_{1 \leq k \leq n} |S_k|\| \leq 2\sqrt{n}\Theta_2$. Since

$$\sum_{i=k+1}^n (X_{i-k} - \bar{X}_n)(X_i - \bar{X}_n) - \sum_{i=k+1}^n X_{i-k}X_i = -\bar{X}_n \sum_{i=1}^{n-k} X_i + \bar{X}_n \sum_{i=1}^k X_i - k\bar{X}_n^2,$$

we have (18). The proof of the $\check{\gamma}_k$ version of Theorem 2 is similar. The assertion on sample autocorrelations can be proved easily, and details are omitted. \square

2.2. Box-Pierce tests

Box-Pierce tests (Box and Pierce, 1970; Ljung and Box, 1978) are commonly used in detecting lack of fit of a particular time series model. After a correct model has been fitted to a set of observations, one would expect the residuals to be close to a sequence of iid random variables, and therefore one should perform some tests for serial correlations as model diagnostics. Suppose $(X_i)_{1 \leq i \leq n}$ is an iid sequence, let \hat{r}_k be its sample autocorrelations. Then the distribution of $Q_n(K) := n \sum_{k=1}^K \hat{r}_K^2$ is approximately χ_K^2 . Logically, it is not sufficient to consider a fixed number of correlations as the number of observations increases, because there may be some dependencies at large lags. We present a normal theory about the Box-Pierce test statistic, which allows the number of correlations included in Q_n to go to infinity.

Theorem 4. *Assume $X_i \in \mathcal{L}^8$, $\mathbb{E}X_i = 0$ and $\sum_{k=0}^{\infty} k^6 \delta_8(k) < \infty$. If $s_n \rightarrow \infty$ and $s_n = O(n^\beta)$ for some $\beta < 1$, then*

$$\frac{1}{\sqrt{s_n}} \sum_{k=1}^{s_n} [n(\hat{\gamma}_k - (1 - k/n)\gamma_k)^2 - (1 - k/n)\sigma_0] \Rightarrow \mathcal{N} \left(0, 2 \sum_{k \in \mathbb{Z}} \sigma_k^2 \right).$$

To see the connection to the Box-Pierce test, we have the following corollary on autocorrelations. Using the same argument, we can show that the same asymptotic law holds for the similar Ljung-Box test statistic $Q_{LB} = n(n+2) \sum_{k=1}^K \hat{r}_K^2 / (n-k)$.

Corollary 5. *Under the conditions of Theorem 4, the same result holds if $\hat{\gamma}_k$ is replaced by $\check{\gamma}_k$. Furthermore,*

$$\frac{1}{\sqrt{s_n}} \sum_{k=1}^{s_n} [n(\hat{r}_k - (1 - k/n)r_k)^2 - (1 - k/n)\sigma_0/\gamma_0^2] \Rightarrow \mathcal{N} \left(0, \frac{2}{\gamma_0^4} \sum_{k \in \mathbb{Z}} \sigma_k^2 \right). \quad (19)$$

Remark 1. Theorem 4 clarifies an important historical issue in testing of correlations. If $\gamma_k = 0$ for all $k \geq 1$, which means X_i are uncorrelated; then $\sigma_0 = \gamma_0^2$ and $\sigma_k = 0$ for all $|k| \geq 1$, and (19) becomes

$$\frac{1}{\sqrt{s_n}} \sum_{k=1}^{s_n} [n\hat{r}_k^2 - (1 - k/n)] \Rightarrow \mathcal{N}(0, 2). \quad (20)$$

In an influential paper, Romano and Thombs (1996) argued that, for fixed K , the chi-squared approximation for $Q_n(K)$ does not hold if X_i are only uncorrelated but not independent. One of the main reasons is that for fixed K , $\hat{r}_1, \dots, \hat{r}_K$ are not asymptotically independent if X_i are not independent. However, interestingly, the situation is different if the number of correlations included in Q_n can increase to infinity. According to (5), $\sqrt{n}\hat{\gamma}_{k_n}$ and $\sqrt{n}\hat{\gamma}_{k_n+h}$ are asymptotically independent if $h > 0$ and $k_n \rightarrow \infty$, because the asymptotic covariance is $\sigma_h = 0$. Therefore, the original Box-Pierce approximation of $Q_n(s_n)$ by $\chi_{s_n}^2$, with unbounded s_n , is still asymptotically valid in the sense of (20) since $(\chi_{s_n}^2 - s_n)/\sqrt{s_n} \Rightarrow \mathcal{N}(0, 2)$ as $s_n \rightarrow \infty$. This observation again suggests that the asymptotic behaviors for bounded and unbounded lags are different. A similar observation has been made in Shao (2011), whose result also suggests that (20) is true under the assumption that $\delta_8(k) = O(\rho^k)$ for some $0 < \rho < 1$. Our condition $\sum_{k=1}^{\infty} k^6 \delta_8(k) < \infty$ is much weaker.

The next theorem consists of two separate but closely related parts, one is on the estimation of $\sigma_0 = \sum_{k \in \mathbb{Z}} \gamma_k^2$, and the other is related to the power of the Box-Pierce test. Define the projection operator

$$\mathcal{P}^j \cdot = \mathbb{E}(\cdot | \mathcal{F}_{-\infty}^j) - \mathbb{E}(\cdot | \mathcal{F}_{-\infty}^{j-1}), \text{ where } \mathcal{F}_i^j = \langle \epsilon_i, \epsilon_{i+1}, \dots, \epsilon_j \rangle, i, j \in \mathbb{Z}.$$

Theorem 6. Assume $X_i \in \mathcal{L}^4$, $\mathbb{E}X_i = 0$ and $\Theta_4 < \infty$. If $s_n \rightarrow \infty$ and $s_n = o(\sqrt{n})$, then

$$\sqrt{n} \left(\sum_{k=-s_n}^{s_n} \hat{\gamma}_k^2 - \sum_{k=-s_n}^{s_n} \gamma_k^2 \right) \Rightarrow \mathcal{N}(0, 4\|D'_0\|^2), \quad (21)$$

where $D'_0 = \sum_{i=0}^{\infty} \mathcal{P}^0(X_i Y_i)$ with $Y_i = \gamma_0 X_i + 2 \sum_{k=1}^{\infty} \gamma_k X_{i-k}$. Furthermore, if $\sum_{k=1}^{\infty} \gamma_k^2 > 0$, then

$$\sqrt{n} \left(\sum_{k=1}^{s_n} \hat{\gamma}_k^2 - \sum_{k=1}^{s_n} \gamma_k^2 \right) \Rightarrow \mathcal{N}(0, 4\|D_0\|^2), \quad (22)$$

where $D_0 = \sum_{i=0}^{\infty} \mathcal{P}^0(X_i Y_i)$ with $Y_i = \sum_{k=1}^{\infty} \gamma_k X_{i-k}$.

Corollary 7. Under conditions of Theorem 6, the same results hold if $\hat{\gamma}_k$ is replaced by $\check{\gamma}_k$. Furthermore, there exist positive numbers τ_1^2 and τ_2^2 such that

$$\sqrt{n} \left(\sum_{k=1}^{s_n} \hat{r}_k^2 - \sum_{k=1}^{s_n} r_k^2 \right) \Rightarrow \mathcal{N}(0, \tau_1^2) \quad \text{and} \quad \sqrt{n} \left(\sum_{k=-s_n}^{s_n} \hat{r}_k^2 - \sum_{k=-s_n}^{s_n} r_k^2 \right) \Rightarrow \mathcal{N}(0, \tau_2^2).$$

As an immediate application, we consider testing whether (X_i) is an uncorrelated sequence. According to (20), we can use the test statistic

$$T_n := \frac{1}{\sqrt{s_n}} \left[Q_n(s_n) - \frac{s_n(2n - s_n - 1)}{2n} \right],$$

whose asymptotic distribution under the null hypothesis is $\mathcal{N}(0, 2)$. The null is rejected when $T_n > \sqrt{2}z_{1-\alpha}$, where $z_{1-\alpha}$ is the $(1-\alpha)$ -th quantile of a standard normal random variable Z . However, under the alternative hypothesis $\sum_{k=1}^{\infty} r_k^2 > 0$, the distribution of T_n should be approximated according to Corollary 7, and the asymptotic power is

$$P\left(T_n > \sqrt{2}z_{1-\alpha}\right) \approx P\left(\tau_1 Z > \frac{\sqrt{2s_n} \cdot z_{1-\alpha}}{\sqrt{n}} + \frac{s_n(2n - s_n - 1)}{2n^{3/2}} - \sqrt{n} \sum_{k=1}^{s_n} r_k^2\right),$$

which increases to 1 as n goes to infinity.

3. A Simulation Study

Suppose $(r_k^{(0)})$ is a sequence of autocorrelations, one might be interested in the hypothesis test that $r_k = r_k^{(0)}$ for all $k \geq 1$. This hypothesis is, however, impossible to test in practice, except in some special parametric cases. A more tractable hypothesis is

$$\mathbf{H}_0 : \quad r_k = r_k^{(0)} \quad \text{for } 1 \leq k \leq s_n. \quad (23)$$

In traditional asymptotic theory, one often assumes that s_n is a fixed constant, for example, the popular Box-Pierce test for serial correlation. Our results in the previous section provide both \mathcal{L}^∞ and \mathcal{L}^2 based tests, which allow s_n to grow as n increases. Nonetheless, the asymptotic tests can perform poorly when the sample size n is not large enough, namely, there may exist noticeable differences between the true and nominal probabilities of rejecting \mathbf{H}_0 (hereafter referred as error in rejection probability or ERP). In a recent paper, Horowitz *et al.* (2006) showed that the Box-Pierce test with bootstrap-based p -values can significantly reduce the ERP. They used the blocks of blocks bootstrapping with overlapping blocks (hereafter referred as BOB) invented by Künsch (1989). For finite sample, our \mathcal{L}^2 based test is similar as the traditional Box-Pierce test considered in their paper, so in this section our focus will be on the \mathcal{L}^∞ based tests. We shall provide simulation evidence showing that the BOB works reasonably well.

Throughout this section, we let the innovations ϵ_i be iid standard normal random variables, and consider the following four models.

$$\text{I.I.D.:} \quad X_i = \epsilon_i \quad (24)$$

$$\text{AR(1):} \quad X_i = bX_{i-1} + \epsilon_i \quad (25)$$

$$\text{Bilinear:} \quad X_i = (a + b\epsilon_i)X_{i-1} + \epsilon_i \quad (26)$$

$$\text{ARCH:} \quad X_i = \sqrt{a + bX_{i-1}^2} \cdot \epsilon_i. \quad (27)$$

We generate each process with length $n = 2 \times 10^7$, and compute

$$a_{2s_n}^{-1} \left(\max_{1 \leq k \leq s_n} \sqrt{n} |\hat{r}_k - (1 - k/n)r_k| / \sqrt{\hat{\sigma}_0} - b_{2s_n} \right) \quad (28)$$

with $s_n = 5 \times 10^5$ and $\hat{\sigma}_0 = \sum_{k=-t_n}^{t_n} \hat{r}_k^2$, where t_n is chosen as $t_n = \lfloor n^{1/3} \rfloor = 271$. Based on 1000 repetitions, we plot the empirical distribution functions in Figure 1. We see that these four empirical curves are close to the one for the Gumbel distribution, which confirms our theoretical results.

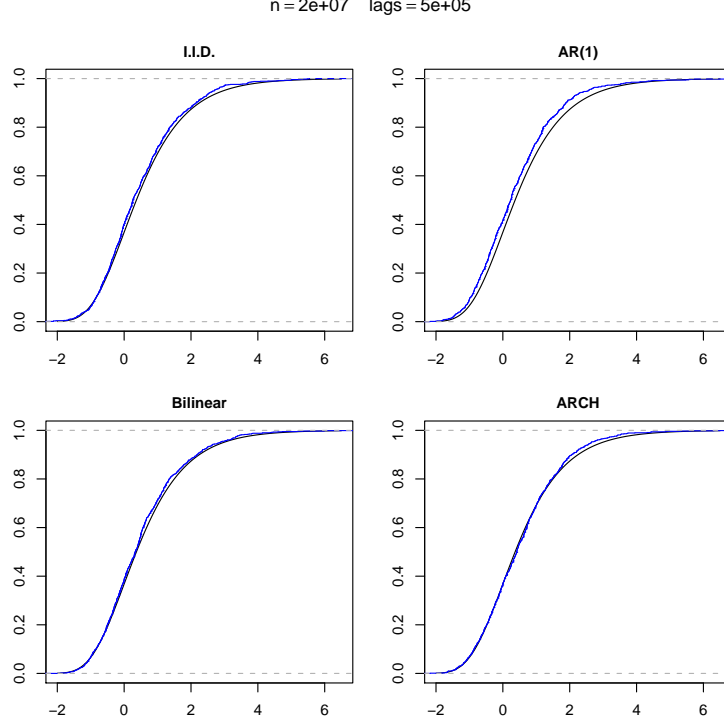


FIG 1. Empirical distribution functions for quantities in (28). We choose $b = 0.5$ for model (25), $a = b = 0.4$ for model (26), and $a = b = 0.25$ for model (27). The black line gives the true distribution function of the Gumbel distribution.

On the other hand, these empirical distributions are not very close to the limiting one if the sample size is not large, because the Gumbel type of convergence in (13) is slow. This is a well-known phenomenon; see for example Hall (1979). It is therefore not reasonable to use the limiting distribution to approximate the finite sample distributions. To perform the test (23), we repeat the BOB procedure as described in Horowitz *et al.* (2006) (called SBOB in their paper). Since in the bootstrapped tests, the test statistics are not to be compared with the limiting distribution, we can ignore the norming constants in (28) and simply use the following test statistics

$$M_n = \max_{1 \leq k \leq s_n} |r_k - (1 - k/n)r_k^{(0)}| \text{ and } \mathcal{M}_n = \frac{M_n}{\sqrt{\hat{\sigma}_0}},$$

where \mathcal{M}_n is the self-normalized version with σ_0 estimated as $\hat{\sigma}_0 = \sum_{k=-t_n}^{t_n} \hat{r}_k^2$, with $t_n = \min\{\lfloor n^{1/3} \rfloor, s_n\}$. For simplicity, we refer these two tests as M -test and \mathcal{M} -test, respectively.

From the series X_1, \dots, X_n , for some specified number of lags s_n that will be included in the test and block size \mathfrak{b}_n , form $Y_i = (X_i, X_{i+1}, \dots, X_{i+s_n})^\top$, $1 \leq i \leq n - s_n$ and blocks $\mathcal{B}_j = (Y_j, Y_{j+1}, \dots, Y_{j+\mathfrak{b}_n-1})$, $1 \leq j \leq n - s_n - \mathfrak{b}_n + 1$. For simplicity assume $h_n = n/\mathfrak{b}_n$ is an integer. Suppose $Y_\#^*$ is obtained by sampling

a block $\mathcal{B}_\#$ from the set of blocks $\{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_{n-s_n-b_n+1}\}$, and then sampling a column from $\mathcal{B}_\#$, let $\text{Cov}_\#$ represent the covariance of the bootstrap distribution of $Y_\#$, conditional on (X_1, X_2, \dots, X_n) . Denote by $Y_\#^j$ the j -th entry of $Y_\#$, set

$$r_k^{(e)} = \frac{\text{Cov}_\#(Y_\#^1, Y_\#^{k+1})}{\sqrt{\text{Cov}_\#(Y_\#^1, Y_\#^1) \cdot \text{Cov}_\#(Y_\#^{k+1}, Y_\#^{k+1})}}.$$

The explicit formula of $r_k^{(e)}$ was also given in Horowitz *et al.* (2006). The BOB algorithm is as follows.

1. Sample h_n times with replacement from $\{\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_{n-s_n-b_n+1}\}$ to obtain blocks $\{\mathcal{B}_1^*, \mathcal{B}_2^*, \dots, \mathcal{B}_{h_n}^*\}$, which are laid end-to-end to form a series of vectors $(Y_1^*, Y_2^*, \dots, Y_n^*)$.
2. Pretend that $(Y_1^*, Y_2^*, \dots, Y_n^*)$ is a random sample of size n from some s_n -dimensional population distribution, let r_k^* be the sample correlation of the first entry and the $(k+1)$ -th entry. Then calculate the test statistic $M_n^* = \max_{1 \leq k \leq s_n} |r_k^* - r_k^{(e)}|$ and $\mathcal{M}_n^* = M_n^* / \sqrt{\sigma_0^*}$, where $\sigma_0^* = \sum_{k=-t_n}^{t_n} (r_k^*)^2$.
3. Repeat steps 1 and 2 for N times. The bootstrap p -value of the M -test is given by $\#(M_n^* > M_n) / N$. For a nominal level α , we reject \mathbf{H}_0 if $\#(M_n^* > M_n) / N < \alpha$. The \mathcal{M} -test is performed in the same manner.

We compare the BOB tests and the asymptotic tests for the four models listed at the beginning of this section, with $a = .4$ for (25), $a = b = .4$ for (26) and $a = b = .25$ for (27). We set the series length as $n = 1800$, and consider four choices of s_n : $\lfloor \log(n) \rfloor = 7$, $\lfloor n^{1/3} \rfloor = 12$, $\lfloor \sqrt{n} \rfloor = 42$ and 25. The BOB tests are performed with $N = 999$, and the asymptotic tests are carried out by comparing $a_{2s_n}^{-1}(\sqrt{n}\mathcal{M}_n - b_{2s_n})$ with the corresponding quantiles of the Gumbel distribution. The empirical rejection probabilities based on 10,000 repetitions are reported in Table 1. All probabilities are given in percentages. For all cases, we see that the asymptotic tests are too conservative, and the ERP are quite large. At the nominal level 1%, the rejection probabilities are often less than or around 0.1%, and at most 0.51%; while at nominal level 10%, they are often less than 3% and at most 6.4%. Except for the bilinear models with $s_n = 7$ and $s_n = 12$, the bootstrapped tests significantly reduce the ERP, which are often less than 0.2% at nominal level 1%, less than .5% at level 5%, and less than 1% at level 10%. The performance of M -test and \mathcal{M} -test are similar, with the former being slightly more conservative. The BOB tests are roughly insensitive to the block size, which provides additional evidence of the findings on BOB tests in Davison and Hinkley (1997).

The bootstrapped tests still perform relatively poorly for bilinear models when s_n is small (7 and 12). This is possibly due to the heavy-tailedness of the bilinear process. Tong (1981) gave necessary conditions for the existence of even order moments. On the other hand, Horowitz *et al.* (2006) showed that the iterated bootstrapping further reduce the ERP. It is of interest to see whether the iterated procedure has the same effect for the \mathcal{L}^∞ based tests, in particular, whether it makes the ERP reasonably small for the bilinear models when s_n is small. The simulation for the iterated bootstrapping will be computationally expensive and we do not pursue it here.

TABLE 1
Empirical rejection probabilities (in percentages)

Test	$s_n = 7$			$s_n = 12$			$s_n = 25$			$s_n = 42$		
	1	5	10	1	5	10	1	5	10	1	5	10
I.I.D.	.00	.34	1.6	.02	.69	2.3	.03	.93	3.2	.04	1.0	3.3
$\mathbf{b}_n = 5$	1.3	5.1	10.0	1.1	5.2	9.8	.95	4.7	9.3	1.0	4.7	9.6
	1.4	5.3	10.4	1.2	5.6	10.5	1.1	5.1	10.1	1.1	5.1	10.2
$\mathbf{b}_n = 10$.83	4.8	10.0	1.1	4.9	9.6	1.1	4.9	10.1	.65	4.3	8.9
	.94	5.1	10.3	1.2	5.4	10.3	1.1	5.5	11.0	.78	4.7	9.6
AR(1)	.01	.17	1.2	.01	.36	1.8	.02	.77	2.5	.02	.88	2.8
$\mathbf{b}_n = 10$	1.3	5.7	10.9	1.3	5.5	11.4	1.3	5.5	10.9	1.1	5.7	11.2
	1.3	5.7	11.2	1.4	5.9	11.7	1.3	6.0	11.5	1.2	6.0	11.7
$\mathbf{b}_n = 20$.98	5.5	10.9	1.0	5.8	11.3	1.1	5.3	10.6	.86	4.9	10.5
	1.0	5.7	11.0	1.1	6.1	11.9	1.2	5.6	11.0	.83	5.0	10.9
Bilinear	.34	2.8	6.4	.43	2.5	5.8	.51	2.5	5.9	.40	2.8	5.9
$\mathbf{b}_n = 10$	2.8	8.7	14.4	1.8	7.1	12.7	1.2	6.1	12.0	1.2	5.4	10.9
	2.7	8.6	14.5	1.8	7.3	12.9	1.3	6.2	12.2	1.1	5.5	11.1
$\mathbf{b}_n = 20$	2.7	8.4	14.6	2.1	7.2	13.5	1.5	6.3	12.0	1.3	5.2	10.8
	2.5	8.3	14.6	2.1	7.5	13.9	1.5	6.2	12.0	1.2	5.3	10.9
ARCH	.05	.82	3.2	.06	1.5	3.9	.09	1.3	4.0	.12	1.4	4.4
$\mathbf{b}_n = 10$.99	5.0	10.5	1.2	4.9	9.7	.80	4.6	9.9	.82	4.7	9.3
	1.1	5.4	10.9	1.4	5.3	10.4	.92	5.1	10.7	.94	5.1	10.2
$\mathbf{b}_n = 20$.86	5.1	10.5	1.0	5.0	10.3	.69	4.8	9.7	.63	4.3	8.9
	.98	5.5	11.0	1.2	5.6	11.0	.89	5.1	10.4	.76	4.7	9.5

The values 1, 5, 10 in the 2nd row indicate nominal levels in percentages. The numbers in the third row starting with the model name ‘‘I.I.D.’’ are for the asymptotic tests. The fourth row starting with $\mathbf{b}_n = 5$ is for BOB M -tests with block size 5. The fifth row is for BOB \mathcal{M} -tests with the same block size 5. Other rows should be read similarly.

4. Proofs

This section provides proofs for the results in Section 2. For readability we list the notation here. For a random variable X , write that $X \in \mathcal{L}^p$, $p > 0$, if $\|X\|_p := (\mathbb{E}|X|^p)^{1/p} < \infty$. Write $\|X\| = \|X\|_2$ if $p = 2$. To express centering of random variables concisely, we define the operator \mathbb{E}_0 as $\mathbb{E}_0 X := X - \mathbb{E}X$. For a vector $\mathbf{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d$, let $|\mathbf{x}|$ be the usual Euclidean norm, $|\mathbf{x}|_\infty := \max_{1 \leq i \leq d} |x_i|$, and $|\mathbf{x}|_\bullet := \min_{1 \leq i \leq d} |x_i|$. For a square matrix A , $\rho(A)$ denotes the operator norm defined by $\rho(A) := \max_{|\mathbf{x}|=1} |A\mathbf{x}|$. Let us make some convention on the constants. We use C , c and \mathcal{C} for constants. The notation \mathcal{C}_p is reserved for the constant appearing in Burkholder’s inequality, see (30). The values of C may vary from place to place, while the value of c is fixed within the statement and the proof of a theorem (or lemma). A constant with a symbolic subscript is used to emphasize the dependence of the value on the subscript.

The framework (9) is particularly suited for two classical tools for dealing with dependent sequences, martingale approximation and m -dependence approximation. For $i \leq j$, define $\mathcal{F}_i^j = \langle \epsilon_i, \epsilon_{i+1}, \dots, \epsilon_j \rangle$ be the σ -field generated by the innovations $\epsilon_i, \epsilon_{i+1}, \dots, \epsilon_j$, and the projection operator $\mathcal{H}_i^j(\cdot) = \mathbb{E}(\cdot | \mathcal{F}_i^j)$. Set $\mathcal{F}_i := \mathcal{F}_i^\infty$, $\mathcal{F}^j := \mathcal{F}_{-\infty}^j$, and define \mathcal{H}_i and \mathcal{H}^j similarly. Define the projection operator $\mathcal{P}^j(\cdot) = \mathcal{H}^j(\cdot) - \mathcal{H}^{j-1}(\cdot)$, and $\mathcal{P}_i(\cdot) = \mathcal{H}_i(\cdot) - \mathcal{H}_{i+1}(\cdot)$, then $(\mathcal{P}^j(\cdot))_{j \in \mathbb{Z}}$ and $(\mathcal{P}_{-i}(\cdot))_{i \in \mathbb{Z}}$ become martingale difference sequences with respect to the filtrations (\mathcal{F}^j) and (\mathcal{F}_{-i}) , respectively. For $m \geq 0$, define $\tilde{X}_i = \mathcal{H}_{i-m} X_i$, then $(\tilde{X}_i)_{i \in \mathbb{Z}}$ is a $(m+1)$ -dependent sequence.

4.1. Some Useful Inequalities

We collect in Proposition 8 some useful facts about physical dependence measures and martingale and m -dependence approximations. We expect that it will be useful in other asymptotic problems that involve sample covariances. Hence for convenience of other researchers, we provide explicit upper bounds.

We now introduce a moment inequality (29) which follows from the Burkholder inequality (see Burkholder, 1988). Let (D_i) be a martingale difference sequence and for every i , $D_i \in \mathcal{L}^p$, $p > 1$, then

$$\|D_1 + D_2 + \dots + D_n\|_p^{p'} \leq C_p^{p'} \left(\|D_1\|_p^{p'} + \|D_2\|_p^{p'} + \dots + \|D_n\|_p^{p'} \right), \quad (29)$$

where $p' = \min\{p, 2\}$, and the constant

$$C_p = (p-1)^{-1} \text{ if } 1 < p < 2 \text{ and } = \sqrt{p-1} \text{ if } p \geq 2. \quad (30)$$

We note that when $p > 2$, the constant C_p in (29) equaled to $p-1$ in Burkholder (1988), and it was improved to $\sqrt{p-1}$ by Rio (2009).

Proposition 8. 1. Assume $\mathbb{E}X_i = 0$ and $p > 1$. Recall that $p' = \min(p, 2)$.

$$\|\mathcal{P}^0 X_i\|_p \leq \delta_p(i) \quad \text{and} \quad \|\mathcal{P}_0 X_i\|_p \leq \delta_p(i) \quad (31)$$

$$\kappa_p := \|X_0\|_p \leq C_p \Psi_p \quad (32)$$

$$\left\| \sum_{i=1}^n c_i X_i \right\|_p \leq C_p A_n \Theta_p, \quad \text{where } A_n = \left(\sum_{i=1}^n |c_i|^{p'} \right)^{1/p'} \quad (33)$$

$$|\gamma_k| \leq \zeta_2(k), \quad \text{where } \zeta_p(k) := \sum_{j=0}^{\infty} \delta_p(j) \delta_p(j+k) \quad (34)$$

$$\left\| \sum_{i=1}^n (X_{i-k} X_i - \gamma_k) \right\|_{p/2} \leq 2C_{p/2} \kappa_p \Theta_p \sqrt{n}, \quad \text{when } p \geq 4 \quad (35)$$

$$\left\| \sum_{i,j=1}^n c_{i,j} (X_i X_j - \gamma_{i-j}) \right\|_{p/2} \leq 4C_{p/2} C_p \Theta_p^2 B_n \sqrt{n}, \quad \text{when } p \geq 4 \quad (36)$$

where $B_n^2 = \max\{\max_{1 \leq i \leq n} \sum_{j=1}^n c_{i,j}^2, \max_{1 \leq j \leq n} \sum_{i=1}^n c_{i,j}^2\}$.

2. For $m \geq 0$, define $\tilde{X}_i = \mathcal{H}_{i-m}X_i$. For $p > 1$, let $\tilde{\delta}_p(\cdot)$ be the physical dependence measures for the sequence (\tilde{X}_i) . Then

$$\tilde{\delta}_p(i) \leq \delta_p(i) \quad (37)$$

$$\|X_0 - \tilde{X}_0\|_p \leq \mathcal{C}_p \Psi_p(m+1) \quad (38)$$

$$\left\| \sum_{i=1}^n c_i(X_i - \tilde{X}_i) \right\|_p \leq \mathcal{C}_p A_n \Theta_p(m+1) \quad (39)$$

$$\left\| \sum_{i=k+1}^n \left(X_{i-k}X_i - \gamma_k - \tilde{X}_{i-k}\tilde{X}_i + \tilde{\gamma}_k \right) \right\|_p \leq 4\mathcal{C}_p(n-k)^{1/p'} \kappa_{2p} \Delta_{2p}(m+1). \quad (40)$$

Proof. The inequalities (31) and (37) are obtained by the first principle. Since $X_{i-k} = \sum_{j \in \mathbb{Z}} \mathcal{P}^j X_{i-k}$ and $X_i = \sum_{j \in \mathbb{Z}} \mathcal{P}^j X_i$, we have

$$|\gamma_k| = \left| \sum_{j=-k}^{\infty} \mathbb{E}[(\mathcal{P}^{-j}X_0)(\mathcal{P}^{-j}X_k)] \right| \leq \delta_2(j)\delta_2(j+k) \leq \zeta_k,$$

which proves (34). For (36), it can be similarly proved as Proposition 1 of Liu and Wu (2010), and (39) was given by Lemma 1 of the same paper. (33) is a special case of (39). Define $Y_i = X_{i-k}X_i$, then (Y_i) is also a stationary process of the form (9). By Hölder's inequality, $\|Y_i - \Omega_0(Y_i)\|_{p/2} \leq 2\kappa_p[\delta_p(i) + \delta_p(i-k)]$. Applying (33) to (Y_i) , we obtain (35). To see (38), we first write $X_m - \tilde{X}_m = \sum_{j=1}^{\infty} \mathcal{P}_{-j}X_m$. Since $\|\mathcal{P}_{-j}X_m\|_p \leq \delta_p(m+j)$, and $(\mathcal{P}_{-j}X_m)_{j \geq 1}$ is a martingale difference sequence, by (29), we have

$$\|X_0 - \tilde{X}_0\|_p^{p'} \leq \mathcal{C}_p^{p'} \sum_{j=1}^{\infty} \|\mathcal{P}_{-j}X_m\|_p^{p'} \leq \mathcal{C}_p^{p'} \sum_{j=1}^{\infty} [\delta_p(m+j)]^{p'} = \mathcal{C}_p^{p'} [\Psi_p(m+1)]^{p'}.$$

The above argument also leads to (32). Using a similar argument as in the proof of Theorem 2 of Wu (2009), we can show (40). Details are omitted. \square

4.2. Proof of Theorem 1

The proof is quite complicated and will be divided into several steps. We first give the outline.

4.1.0. Outline

Step 1: m -dependence approximation. Define $R_{n,k} = \sum_{i=k+1}^n (X_{i-k}X_i - \gamma_k)$. Set $m_n = \lfloor n^\beta \rfloor$, $0 < \beta < 1$. Define $\tilde{X}_i = \mathcal{H}_{i-m_n}X_i$, $\tilde{\gamma}_k = \mathbb{E}(\tilde{X}_0\tilde{X}_k)$, and $\tilde{R}_{n,k} = \sum_{i=k+1}^n (\tilde{X}_{i-k}\tilde{X}_i - \tilde{\gamma}_k)$. We next show that it suffices to consider $\tilde{R}_{n,k}$.

Lemma 9. Assume $\mathbb{E}X_i = 0$, $X_i \in \mathcal{L}^p$, and $\Theta_p(m) = O(m^{-\alpha})$ for some $p > 4$ and $\alpha > 0$. If $s_n = O(n^\eta)$ with $0 < \eta < \alpha p/2$, then there exists a β such that $\eta < \beta < 1$ and

$$\max_{1 \leq k \leq s_n} |R_{n,k} - \tilde{R}_{n,k}| = o_P\left(\sqrt{n/\log s_n}\right).$$

Step 2: Throw out small blocks. Let $l_n = \lfloor n^\gamma \rfloor$, where $\gamma \in (\beta, 1)$. For each $t_n < k \leq s_n$, we apply the blocking technique and split the integer interval $[k+1, n]$ into alternating large and small blocks

$$\begin{aligned} K_1 &= [k+1, s_n] \\ H_j &= [s_n + (j-1)(2m_n + l_n) + 1, s_n + (j-1)(2m_n + l_n) + l_n]; \quad 1 \leq j \leq w_n - 1, \\ K_{j+1} &= [s_n + (j-1)(2m_n + l_n) + l_n + 1, s_n + j(2m_n + l_n)]; \quad 1 \leq j \leq w_n - 1; \quad \text{and} \\ H_{w_n} &= [s_n + (w_n - 1)(2m_n + l_n) + 1, n], \end{aligned} \tag{41}$$

where w_n is the largest integer such that $s_n + (w_n - 1)(2m_n + l_n) + l_n \leq n$. Denote by $|H|$ the size of a block H . By definition we know $l_n \leq |H_{w_n}| \leq 3l_n$ when n is large enough. For $1 \leq j \leq w_n$ define

$$V_{k,j} = \sum_{i \in K_j, i > k} (\tilde{X}_{i-k} \tilde{X}_i - \tilde{\gamma}_k) \quad \text{and} \quad U_{k,j} = \sum_{i \in H_j} (\tilde{X}_{i-k} \tilde{X}_i - \tilde{\gamma}_k).$$

Note that $w_n \sim n/(2m_n + l_n) \sim n^{1-\gamma}$. We show that the sums over small blocks are negligible.

Lemma 10. *Assume the conditions of Theorem 1. Then*

$$\max_{1 \leq k \leq s_n} \left| \sum_{j=1}^{w_n} V_{k,j} \right| = o_P \left(\sqrt{\frac{n}{\log s_n}} \right).$$

Step 3: Truncate sums over large blocks. We show that it suffices to consider

$$\mathcal{R}_{n,k} = \sum_{j=1}^{w_n} \bar{U}_{k,j}, \quad \text{where} \quad \bar{U}_{k,j} = \mathbb{E}_0 (U_{k,j} I\{|U_{k,j}| \leq \sqrt{n}/(\log s_n)^3\}).$$

Lemma 11. *Assume the conditions of Theorem 1. Then*

$$\max_{1 \leq k \leq s_n} \left| \sum_{j=1}^{w_n} (U_{k,j} - \bar{U}_{k,j}) \right| = o_P \left(\sqrt{\frac{n}{\log s_n}} \right).$$

Step 4: Compare covariance structures. In order to prove Lemma 14, we need the autocovariance structure of $(\mathcal{R}_{n,k}/\sqrt{n})$ to be close to that of (G_k) . However, this only happens when k is large. We show that there exists an $0 < \iota < 1$ such that for $t_n = 3\lfloor s_n^\iota \rfloor$, (i) $\max_{1 \leq k \leq t_n} |\mathcal{R}_{n,k}/\sqrt{n}|$ does not contribute to the asymptotic distribution; and (ii) the autocovariance structure of $(\mathcal{R}_{n,k}/\sqrt{n})$ converges to that of (G_k) uniformly on $t_n < k \leq s_n$.

Lemma 12. *Under conditions of Theorem 1, there exists a constant $0 < \iota < 1$ such that for $t_n = 3\lfloor s_n^\iota \rfloor$,*

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq k \leq t_n} |\mathcal{R}_{n,k}| > \sqrt{\sigma_0 n \log s_n} \right) = 0. \tag{42}$$

Lemma 13. *Let conditions of Theorem 1 be satisfied. Recall that $t_n = 3\lfloor s_n^\iota \rfloor$ from Lemma 12. There exist constants $C_p > 0$ and $0 < \ell < 1$ such that for any $t_n < k \leq k+h \leq s_n$,*

$$|\text{Cov}(\mathcal{R}_{n,k}, \mathcal{R}_{n,k+h})/n - \sigma_h| \leq C_p s_n^{-\ell}.$$

Step 5: Moderate deviations. Let $t_n = 3\lfloor s_n^\ell \rfloor$ be as in Lemma 12. For $t_n < k_1 < k_2 < \dots < k_d \leq s_n$, define $\mathbf{R}_n = (\mathcal{R}_{n,k_1}, \mathcal{R}_{n,k_2}, \dots, \mathcal{R}_{n,k_d})^\top$ and $\mathbf{V} = (G_{k_1}, G_{k_2}, \dots, G_{k_d})^\top$, where (G_k) is defined in (6). Let $\Sigma_n = \text{Cov}(\mathbf{R}_n)$ and $\Sigma = \text{Cov}(\mathbf{V})$. For fixed $x \in \mathbb{R}$, set $z_n = a_{2s_n}x + b_{2s_n}$, where the constants a_n and b_n are defined in (7). In the following lemma we provide a moderate deviation result for \mathbf{R}_n .

Lemma 14. *Assume conditions of Theorem 1. Then there exists a constant $C_{p,d} > 1$ such that for all $t_n < k_1 < k_2 < \dots < k_d \leq s_n$,*

$$|P(|\mathbf{R}_n/\sqrt{n}|_\bullet \geq z_n) - P(|\mathbf{V}|_\bullet \geq z_n)| \leq C_{p,d} \frac{P(|\mathbf{V}|_\bullet \geq z_n)}{(\log s_n)^{1/2}} + C_{p,d} \exp\left\{-\frac{(\log s_n)^2}{C_{p,d}}\right\}.$$

4.2.1. Step 1: m -dependence approximation

Proof of Lemma 9. Recall that $m_n = \lfloor n^\beta \rfloor$ with $\eta < \beta < 1$. We claim

$$\|R_{n,k} - \tilde{R}_{n,k}\|_{p/2} \leq 6C_{p/2}\Theta_p\Theta_p(m_n - k + 1) \cdot \sqrt{n}. \quad (43)$$

It follows that for any $\lambda > 0$

$$\begin{aligned} P\left(\max_{1 \leq k \leq s_n} |R_{n,k} - \tilde{R}_{n,k}| > \lambda \sqrt{n/\log s_n}\right) &\leq \frac{(\log s_n)^{p/4}}{n^{p/4}\lambda^{p/2}} \sum_{k=1}^{s_n} \|R_{n,k} - \tilde{R}_{n,k}\|_{p/2}^{p/2} \\ &\leq C_p \lambda^{-p/2} s_n (\log s_n)^{p/4} n^{-\alpha\beta p/2} \leq C_p \lambda^{-p/2} n^{\eta - \alpha\beta p/2} (\log n)^{p/4}. \end{aligned}$$

Therefore, if $\alpha p/2 > \eta$, then there exists a β such that $\eta < \beta < 1$ and $\eta - \alpha\beta p/2 < 0$, and hence the preceding probability goes to zero as $n \rightarrow \infty$. The proof of Lemma 9 is complete.

We now prove claim (43). For each $1 \leq k \leq s_n$, we have

$$\begin{aligned} \|R_{n,k} - \tilde{R}_{n,k}\|_{p/2} &\leq \left\| \sum_{i=k+1}^n (X_{i-k} - \tilde{X}_{i-k}) \tilde{X}_i \right\|_{p/2} + \left\| \sum_{i=k+1}^n (\mathcal{H}_{i-m_n} X_{i-k})(X_i - \tilde{X}_i) \right\|_{p/2} \\ &\quad + \left\| \sum_{i=k+1}^n \mathbb{E}_0 \left[(X_{i-k} - \mathcal{H}_{i-m_n} X_{i-k})(X_i - \tilde{X}_i) \right] \right\|_{p/2} \end{aligned}$$

Observe that $(\tilde{X}_i \mathcal{P}_{i-k-j} X_{i-k})_{1 \leq i \leq n}$ is a backward martingale difference sequence with respect to \mathcal{F}_{i-k-j} if $j > m_n$, so by the inequality (29),

$$\begin{aligned} \left\| \sum_{i=k+1}^n (X_{i-k} - \tilde{X}_{i-k}) \tilde{X}_i \right\|_{p/2} &\leq \sum_{j=m+1}^{\infty} \left\| \sum_{i=k+1}^n \tilde{X}_i \mathcal{P}_{i-k-j} X_{i-k} \right\|_{p/2} \\ &\leq \sum_{j=m+1}^{\infty} \sqrt{n} C_{p/2} \|\tilde{X}_{j+k} \mathcal{P}_0 X_j\|_{p/2} \\ &\leq C_{p/2} \Theta_p \Theta_p(m_n + 1) \cdot \sqrt{n}. \end{aligned}$$

Similarly we have $\|\sum_{i=k+1}^n (\mathcal{H}_{i-m_n} X_{i-k})(X_i - \tilde{X}_i)\|_{p/2} \leq \sqrt{n} C_{p/2} \Theta_p \Theta_p(m_n + 1)$. Similarly as (39), we get $\|\tilde{X}_{i-k} - \mathcal{H}_{i-m_n} X_{i-k}\|_p \leq \Theta_p(m_n - k + 1)$. Let $Y_{n,i} := (X_{i-k} - \mathcal{H}_{i-m_n} X_{i-k})(X_i - \tilde{X}_i)$. Then

$$\|Y_{n,i} - \Omega_0(Y_{n,i})\|_{p/2} \leq 2[\delta_p(i)\Theta_p(m_n - k + 1) + \delta_p(i - k)\Theta_p(m_n + 1)].$$

Therefore, by (33), it follows that

$$\left\| \sum_{i=k+1}^n \mathbb{E}_0 \left[(X_{i-k} - \mathcal{H}_{i-m_n} X_{i-k})(X_i - \tilde{X}_i) \right] \right\|_{p/2} \leq 4C_{p/2} \Theta_p \Theta_p (m_n - k + 1) \cdot \sqrt{n},$$

and the proof of (43) is complete. \square

4.2.2. Step 2: Throw out small blocks

In this section, as well as many other places in this article, we often need to split an integer interval $[s, t] = \{s, s+1, \dots, t\} \subset \mathbb{N}$ into consecutive blocks $\mathcal{B}_1, \dots, \mathcal{B}_w$ with the size m . Since $s - t + 1$ may not be a multiple of m , we make the convention that unless the size of the last block is specified clearly, it has the size $m \leq |\mathcal{B}_w| < 2m$, and all the other ones have the same size m .

Proof of Lemma 10. It suffices to show that for any $\lambda > 0$,

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{s_n} P \left(\left| \sum_{j=1}^{w_n} V_{k,j} \right| \geq \lambda \sqrt{\frac{n}{\log s_n}} \right) = 0.$$

Observe that $V_{k,j}, 1 \leq j \leq w_n$, are independent. By (35), $\|V_{k,j}\| \leq 2|K_j|^{1/2} \kappa_4 \Theta_4$. By Corollary 1.6 of Nagaev (1979), for any $M > 1$, there exists a constant $C_M > 1$ such that

$$\begin{aligned} P \left(\left| \sum_{j=1}^{w_n} V_{k,j} \right| \geq \lambda \sqrt{\frac{n}{\log s_n}} \right) &\leq \sum_{j=1}^{w_n} P \left(|V_{k,j}| \geq C_M^{-1} \lambda \sqrt{n/\log s_n} \right) + \left(\frac{4e^2 \kappa_4^2 \Theta_4^2 \sum_{j=1}^{w_n} |K_j|}{C_M^{-1} \lambda^2 n / \log s_n} \right)^{C_M/2} \\ &\leq \sum_{j=1}^{w_n} P \left(|V_{k,j}| \geq C_M^{-1} \lambda \sqrt{n/\log n} \right) + C_M (n^{\beta-\gamma} \log n)^{C_M/2} \\ &\leq \sum_{j=1}^{w_n} P \left(|V_{k,j}| \geq C_M^{-1} \sqrt{n/\log n} \right) + n^{-M}. \end{aligned} \quad (44)$$

where we resolve the constant λ into the constant C_M in the last inequality. It remains to show that

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{s_n} \sum_{j=1}^{w_n} P(|V_{k,j}| \geq q_1 \delta \phi_n) = 0, \text{ where } \phi_n = \sqrt{\frac{n}{\log n}}, \quad (45)$$

holds for any $\delta > 0$, where q_1 is the smallest integer such that $\beta^{q_1} < \min\{(p-4)/p, (p-2-2\eta)/(p-2)\}$. This choice of q_1 will be explained later. We adopt the technique of successive m -dependence approximations from Liu and Wu (2010) to prove (45).

For $q \geq 1$, set $m_{n,q} = \lfloor n^{\beta^q} \rfloor$. Define $X_{i,q} = \mathcal{H}_{i-m_{n,q}} X_i$, $\gamma_{k,q} = \mathbb{E}(X_{0,q} X_{k,q})$, and

$$V_{k,j,q} = \sum_{i \in K_j, i > k} (X_{i-k,q} X_{i,q} - \gamma_{k,q}).$$

In particular, $m_{n,1}$ is same as m_n defined in Step 2, and $V_{k,j,1} = V_{k,j}$. Without loss of generality assume $s_n \leq \lfloor n^\eta \rfloor$. Let q_0 be such that $\beta^{q_0+1} \leq \eta < \beta^{q_0}$. We first consider the difference between $V_{k,j,q}$ and $V_{k,j,q+1}$

for $1 \leq q < q_0$. Split the block K_j into consecutive small blocks $\mathcal{B}_1, \dots, \mathcal{B}_{w_{n,q}}$ with size $2m_{n,q}$. Define

$$V_{k,j,q,t}^{(0)} = \sum_{i \in \mathcal{B}_t} (X_{i-k,q} X_{i,q} - \gamma_{k,q}) \quad \text{and} \quad V_{k,j,q,t}^{(1)} = \sum_{i \in \mathcal{B}_t} (X_{i-k,q+1} X_{i,q+1} - \gamma_{k,q+1}). \quad (46)$$

Observe that $V_{k,j,q,t_1}^{(0)}$ and $V_{k,j,q,t_2}^{(0)}$ are independent if $|t_1 - t_2| > 1$. Similar as (44), for any $M > 1$, there exists a constant $C_M > 1$ such that, for sufficiently large n ,

$$\begin{aligned} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) &= P\left[\left|\sum_{t=1}^{w_{n,q}} \left(V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)}\right)\right| \geq \delta \phi_n\right] \\ &\leq \sum_{t=1}^{w_{n,q}} P\left(\left|V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)}\right| \geq C_M^{-1} \phi_n\right) + n^{-M}. \end{aligned} \quad (47)$$

Similarly as (43), we have $\left\|V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)}\right\|_{p/2} \leq C_p |\mathcal{B}_t|^{1/2} m_{n,q+1}^{-\alpha}$. It follows that

$$\begin{aligned} \sum_{k=1}^{s_n} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) &\leq C_{p,M} n^\eta n^{1-\gamma} \left(n^{-M} + \frac{n^\gamma m_{n,q}^{p/4} m_{n,q+1}^{-\alpha p/2}}{m_{n,q} (n/\log n)^{p/4}}\right) \\ &\leq C_{p,M} \left(n^{\eta+1-\gamma-M} + n^\eta n^{1-p/4} m_{n,q}^{p/4-1-\alpha \beta p/2}\right). \end{aligned}$$

Under the condition (16), there exists a $0 < \beta < 1$, such that

$$\sum_{k=1}^{s_n} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) \leq C_{p,M} \left(n^{\eta+1-\gamma-M} + n^{\eta+1-p/4+\beta q(p/4-1-\alpha \beta p/2)}\right) \rightarrow 0.$$

Recall that q_1 is the smallest integer such that $\beta^{q_1} < \min\{(p-4)/p, (p-2-2\eta)/(p-2)\}$. We now consider the difference between $V_{k,j,q}$ and $V_{k,j,q+1}$ for $q_0 \leq q < q_1$. The problem is more complicated than the preceding case $1 \leq q < q_0$, since now it is possible that $m_{n,q} < k$ for some $1 \leq k \leq s_n$. We consider three cases.

Case 1: $k \geq 2m_{n,q}$. Partition the block K_j into consecutive smaller blocks $\mathcal{B}_1, \dots, \mathcal{B}_{w_{n,q}}$ with same size $m_{n,q}$. Define $V_{k,j,q,t}^{(0)}$ and $V_{k,j,q,t}^{(1)}$ as in (46). Observe that $\left(V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)}\right)_{t \text{ is odd}}$ is a martingale difference sequence with respect to the filtration $(\xi_t := \langle \epsilon_l : l \leq \max\{\mathcal{B}_t\} \rangle)_{t \text{ is odd}}$, and so is the sequence and filtration labelled by even t . Set $\xi_0 = \langle \epsilon_l : l < \min\{\mathcal{B}_1\} \rangle$ and $\xi_{-1} = \langle \epsilon_l : l < \min\{\mathcal{B}_1\} - m_{n,q} \rangle$. For each $1 \leq t \leq w_{n,q}$, define

$$\nu_t^{(l)} = \mathbb{E} \left[\left(V_{k,j,q,t}^{(l)} \right)^2 \middle| \xi_{t-2} \right] = \sum_{i_1, i_2 \in \mathcal{B}_t} X_{i_1-k,q+l} X_{i_2-k,q+l} \gamma_{i_1-i_2,q+l}$$

for $l = 0, 1$. By Lemma 1 of Haeusler (1984), for any $M > 1$, there exists a constant $C_M > 1$ such that

$$\begin{aligned} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) &\leq \sum_{t=1}^{w_{n,q}} P\left(\left|V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)}\right| \geq \sqrt{\frac{n}{(\log n)^3}}\right) + n^{-M} \\ &+ \sum_{l=0,1} 2 \left\{ P\left[\sum_{t \text{ is odd}} \nu_t^{(l)} \geq \frac{C_M^{-1} n}{(\log n)^2}\right] + P\left[\sum_{t \text{ is even}} \nu_t^{(l)} \geq \frac{C_M^{-1} n}{(\log n)^2}\right] \right\}. \end{aligned} \quad (48)$$

By (34), $\sum_{k \in \mathbb{Z}} |\gamma_{k,q+l}|^2 \leq \Theta_2^2$, and hence by (36), $\|\mathcal{V}_t^{(l)}\|_{p/2} \leq C_p m_{n,q}^{1/2}$. Observe that $\mathcal{V}_{t_1}^{(0)}$ and $\mathcal{V}_{t_1}^{(0)}$ are independent if $|t_1 - t_2| > 1$, so similarly as (44), we have

$$\begin{aligned} P \left[\sum_{t \text{ is odd}} \mathcal{V}_t^{(l)} \geq \frac{C_M^{-1} n}{(\log n)^2} \right] &\leq n^{-M} + \sum_{t \text{ is odd}} P \left[\mathcal{V}_t^{(l)} \geq \frac{C_M^{-2} n}{(\log n)^2} \right] \\ &\leq n^{-M} + C_{p,M} \cdot w_{n,q} \cdot n^{-p/2} (\log n)^p \cdot m_{n,q}^{p/4}. \end{aligned}$$

The same inequality holds for the sum over even t . For the first term in (48), we claim that

$$\left\| V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)} \right\|_p \leq C_p \cdot m_{n,q}^{1/2} \cdot m_{n,q+1}^{-\alpha}, \quad (49)$$

which together with the preceding two inequalities implies that

$$P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) \leq C_{p,M} w_{n,q} \cdot n^{-p/2} (\log n)^{3p/2} \left(m_{n,q}^{p/2} \cdot m_{n,q+1}^{-\alpha p} + m_{n,q}^{p/4} \right) + n^{-M}.$$

It follows that under condition (16), there exists a $0 < \beta < 1$ such that

$$\begin{aligned} \sum_{k=2m_{n,q}}^{s_n} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) \\ \leq n^{1+\eta-M} + C_{p,M} \cdot n^{1+\eta-p/2} (\log n)^{3p/2} \left[n^{\beta q(p/2-1-\alpha\beta p)} + n^{\beta q(p/4-1)} \right] = o(1). \end{aligned} \quad (50)$$

Case 2: $k \leq m_{n,q+1}/2$. Partition the block K_j into consecutive smaller blocks $\mathcal{B}_1, \dots, \mathcal{B}_{w_{n,q}}$ with size $3m_{n,q}$. Define $V_{k,j,q,t}^{(0)}$ and $V_{k,j,q,t}^{(1)}$ as in (46). Similarly as (43), we have

$$\left\| V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)} \right\|_{p/2} \leq C_p \cdot m_{n,q}^{1/2} \cdot m_{n,q+1}^{-\alpha}.$$

Similar as (47), for any $M > 1$, there exist a constant $C_M > 1$ such that

$$\begin{aligned} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) &\leq \sum_{t=1}^{w_{n,q}} P \left(\left| V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)} \right| \geq C_M^{-1} \phi_n \right) + n^{-M} \\ &\leq n^{-M} + C_{p,M} \cdot w_{n,q} \cdot n^{-p/4} (\log n)^{p/4} \cdot m_{n,q}^{p/4} \cdot m_{n,q+1}^{-\alpha\beta p/2}. \end{aligned}$$

It follows that that under condition (16), there exists a $0 < \beta < 1$ such that

$$\begin{aligned} \sum_{k=1}^{m_{n,q+1}/2} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) \\ \leq n^{1+\eta-M} + C_{p,M} \cdot n^{1-p/4} (\log n)^{p/4} \cdot \left(n^{\beta q} \right)^{p/4-\alpha\beta p/2} = o(1). \end{aligned} \quad (51)$$

Case 3: $m_{n,q+1}/2 < k < 2m_{n,q}$. We use the same argument as in Case 2. But this time we claim that

$$\left\| V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)} \right\|_{p/2} \leq C_p \left[m_{n,q}^{1/2} \cdot m_{n,q+1}^{-\alpha} + m_{n,q} \zeta_p(k) \right], \quad (52)$$

where $\zeta_p(k)$ is defined in (34). Since $\sum_{k=m}^{\infty} [\zeta_p(k)]^{p/2} \leq [\sum_{k=m}^{\infty} \zeta_p(k)]^{p/2} = O(m^{-\alpha p/2})$, under condition (12), there exist constants $C_{p,M} > 1$ and $0 < \beta < 1$ such that for M large enough

$$\begin{aligned} & \sum_{k > m_{n,q+1}/2}^{2m_{n,q}-1} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) \leq C_{p,M} \cdot n^{1-p/4} (\log n)^{p/4} m_{n,q}^{p/4 - \alpha \beta p/2} \\ & \quad + n^{1+\eta-M} + C_{p,M} \cdot n^{1-p/4} (\log n)^{p/4} \cdot m_{n,q}^{p/2-1} \sum_{k > m_{n,q+1}/2}^{2m_{n,q}-1} [\zeta_p(k)]^{p/2} \\ & \leq n^{1+\eta-M} + C_{p,M} \cdot n^{1-p/4} (\log n)^{p/4} \cdot m_{n,q}^{p/2-1-\alpha \beta p/2} = o(1). \end{aligned} \quad (53)$$

Alternatively, if we use the bound from (40), $\|V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)}\|_{p/2} \leq C_p m_{n,q}^{1/2} \cdot m_{n,q+1}^{-\alpha'}$, it is still true that under condition (12), there exist constants $C_{p,M} > 1$ and $0 < \beta < 1$ such that for M large enough

$$\begin{aligned} & \sum_{k > m_{n,q+1}/2}^{2m_{n,q}-1} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) \\ & \leq n^{1+\eta-M} + C_{p,M} \cdot n^{1-p/4} (\log n)^{p/4} \cdot m_{n,q}^{p/2-1-\alpha' \beta p/2} = o(1). \end{aligned} \quad (54)$$

Combine (50), (51), (53) and (54), we have shown that

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{s_n} \sum_{j=1}^{w_n} P(|V_{k,j,q} - V_{k,j,q+1}| \geq \delta \phi_n) = 0. \quad (55)$$

for $1 \leq q < q_1$. Therefore, to prove (45), it suffices to show

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{s_n} \sum_{j=1}^{w_n} P(|V_{k,j,q_1}| \geq \delta \phi_n) = 0 \quad (56)$$

By considering two cases (i) $2m_{n,q_1} \leq k \leq s_n$ and (ii) $1 \leq k < 2m_{n,q_1}$ under the condition $\beta^{q_1} < \min\{(p-4)/p, (p-2-2\eta)/(p-2)\}$, and using similar arguments as those in proving (55), we can obtain (56). The proof of Lemma 10 is complete.

We now turn to the proof of the two claims (49) and (52). For (52), we have

$$\begin{aligned} \left\| V_{k,j,q,t}^{(0)} - V_{k,j,q,t}^{(1)} \right\|_{p/2} & \leq \left\| \sum_{i \in \mathcal{B}_t} (X_{i-k,q} - X_{i-k,q+1}) X_{i,q+1} \right\|_{p/2} + \left\| \sum_{i \in \mathcal{B}_t} \mathbb{E}_0 [X_{i-k,q+1} (X_{i,q} - X_{i,q+1})] \right\|_{p/2} \\ & \quad + \left\| \sum_{i \in \mathcal{B}_t} \mathbb{E}_0 [(X_{i-k,q} - X_{i-k,q+1}) (X_{i,q} - X_{i,q+1})] \right\|_{p/2} =: I + II + III. \end{aligned}$$

Similarly as in the proof of (43), we have

$$I \leq \mathcal{C}_{p/2} \Theta_p \Theta_p(m_{n,q+1} + 1) \cdot \sqrt{3m_{n,q}} \quad \text{and} \quad III \leq 4 \mathcal{C}_{p/2} \Theta_p \Theta_p(m_{n,q+1} + 1) \cdot \sqrt{3m_{n,q}}.$$

For the second term II , write

$$\mathbb{E}_0 [X_{i-k,q+1} (X_{i,q} - X_{i,q+1})] = \sum_{l_1=0}^{m_{n,q+1}} \sum_{l_2=m_{n,q+1}+1}^{m_{n,q}} \mathbb{E}_0 [(\mathcal{P}_{i-k-l_1} X_{i-k}) (\mathcal{P}_{i-l_2} X_i)].$$

For a pair (l_1, l_2) such that $i - k - l_1 \neq i - l_2$, by the inequality (29), we have

$$\left\| \sum_{i \in \mathcal{B}_t} (\mathcal{P}_{i-k-l_1} X_{i-k}) (\mathcal{P}_{i-l_2} X_i) \right\|_{p/2} \leq \mathcal{C}_{p/2} \delta_p(l_1) \delta_p(l_2) \cdot \sqrt{3m_{n,q}}.$$

For the pairs (l_1, l_2) such that $i - k - l_1 = i - l_2$, by the triangle inequality

$$\left\| \sum_{i \in \mathcal{B}_t} \sum_{l=0}^{m_{n,q}+1} \mathbb{E}_0 [(\mathcal{P}_{i-k-l} X_{i-k}) (\mathcal{P}_{i-k-l} X_i)] \right\|_{p/2} \leq 3m_{n,q} \cdot 2 \sum_{l=0}^{m_{n,q}+1} \delta_p(l) \delta_p(k+l) \leq 6m_{n,q} \zeta_p(k).$$

Putting these pieces together, the proof of (52) is complete. The key observation in proving (49) is that since $k \geq 2m_{n,q}$, $X_{i-k,q}$ and $X_{i,q}$ are independent, hence the product $X_{i-k,q} X_{i,q}$ has finite p -th moment. The rest of the proof is similar to that of (52). Details are omitted. \square

Remark 2. Condition (12) is only used to deal with Case 3, while (16) suffices for the rest of the proof. In fact, for linear processes, one can show that the term $m_{n,q} \zeta_p(k)$ in (52) can be removed, so we have (53) under condition (16) and do not need (54). So (16) suffices for Theorem 1. Furthermore, for nonlinear processes with $\delta_p(k) = O[k^{-(1/2+\alpha)}]$, the term $m_{n,q} \zeta_p(k)$ can also be removed from (52). Details are omitted.

4.2.3. Step 3: Truncate sums over large blocks

Proof of Lemma 11. We need to show for any $\lambda > 0$

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{s_n} P \left(\left| \sum_{j=1}^{w_n} (U_{k,j} - \bar{U}_{k,j}) \right| \geq \lambda \sqrt{\frac{n}{\log s_n}} \right) = 0.$$

Using (35), elementary calculation gives

$$\left\| \tilde{U}_{k,j} - \bar{U}_{k,j} \right\|^2 \leq \frac{\mathbb{E} |\tilde{U}_{k,j}|^{p/2}}{(\sqrt{n}/\log s_n)^{p/2-2}} \leq \frac{(2\mathcal{C}_{p/2} \kappa_p \Theta_p)^{p/2} |H_j|^{p/4} (\log s_n)^{3(p-4)/2}}{n^{(p-4)/4}}. \quad (57)$$

Similarly as (44), for any $M > 1$, there exists a constant $C_M > 1$ such that

$$\begin{aligned} P \left(\left| \sum_{j=1}^{w_n} (U_{k,j} - \bar{U}_{k,j}) \right| \geq \lambda \sqrt{\frac{n}{\log s_n}} \right) &\leq \sum_{j=1}^{w_n} P \left(|U_{k,j} - \bar{U}_{k,j}| \geq C_M^{-1} \lambda \sqrt{\frac{n}{\log s_n}} \right) \\ &\quad + \left(\frac{C_p \sum_{j=1}^{w_n} |H_j|^{p/4} (\log n)^{3p/2}}{C_M^{-1} \lambda^2 n^{p/4}} \right)^{C_M/2} \\ &\leq \sum_{j=1}^{w_n} P \left(|U_{k,j} - \bar{U}_{k,j}| \geq C_M^{-1} \sqrt{\frac{n}{\log s_n}} \right) + n^{-M}. \end{aligned}$$

Therefore, it suffices to show that for any $\delta > 0$,

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{s_n} \sum_{j=1}^{w_n} P \left(|U_{k,j} - \bar{U}_{k,j}| \geq \delta \sqrt{\frac{n}{\log n}} \right) = 0.$$

Since we can use the same arguments as those for (45), Lemma 11 follows. \square

4.2.4. Step 4: Compare covariance structures

Proof of Lemma 12. Since $|\bar{U}_{k,j}| \leq 2\sqrt{n}/(\log s_n)$ and $\mathbb{E}\bar{U}_{k,j}^2 \leq \mathbb{E}U_{k,j}^2 \leq 4(\kappa_4\Theta_4)^2|H_j|$, by Bernstein's inequality (cf. Fact 2.3, Einmahl and Mason, 1997), we have

$$P\left(|\mathcal{R}_{n,k}| > \sqrt{\sigma_0 n \log s_n}\right) \leq \exp\left\{-\frac{(\sigma_0 n \log s_n)/2}{4(\kappa_4\Theta_4)^2 n + n\sqrt{\sigma_0/(\log s_n)}}\right\}.$$

Therefore, for any $0 < \iota < \sigma_0/[8(\kappa_4\Theta_4)^2]$, (42) holds. \square

Proof of Lemma 13. For $1 \leq j \leq w_n$, by (57), we have

$$\begin{aligned} \left|\mathbb{E}(\bar{U}_{k,j}\bar{U}_{k+h,j}) - \mathbb{E}(\tilde{U}_{k,j}\tilde{U}_{k+h,j})\right| &\leq \|\bar{U}_{k,j} - \tilde{U}_{k,j}\| \|\bar{U}_{k+h,j}\| + \|\tilde{U}_{k,j}\| \|\bar{U}_{k+h,j} - \tilde{U}_{k+h,j}\| \\ &\leq 4\kappa_4\Theta_4|H_j|^{1/2} \frac{(2\mathcal{C}_{p/2}\kappa_p\Theta_p)^{p/4}|H_j|^{p/8}(\log s_n)^{3(p-4)/4}}{n^{(p-4)/8}} \\ &\leq C_p|H_j|n^{-(1-\gamma)(p-4)/8}(\log n)^{3(p-4)/4}. \end{aligned}$$

Let $S_{k,j} = \sum_{i \in H_j} (X_{i-k}X_i - \gamma_k)$, by (35) and (43), we have

$$\begin{aligned} \left|\mathbb{E}(S_{k,j}S_{k+h,j}) - \mathbb{E}(\tilde{U}_{k,j}\tilde{U}_{k+h,j})\right| &\leq \|S_{k,j} - \tilde{U}_{k,j}\| \|S_{k+h,j}\| + \|\tilde{U}_{k,j}\| \|S_{k+h,j} - \tilde{U}_{k+h,j}\| \\ &\leq 4\kappa_4\Theta_4|H_j|^{1/2} \cdot 6\Theta_4\Theta_4(m_n - k + 1)|H_j|^{1/2} \leq C|H_j|n^{-\alpha\beta}. \end{aligned}$$

Since $\Theta_4(m) = O(m^{-\alpha})$, elementary calculation shows that $\Delta_4(m) = O(n^{-\alpha^2/(1+\alpha)})$, which together with Lemma 24 implies that if $k > t_n$,

$$\begin{aligned} \left|\mathbb{E}(\tilde{U}_{k,j}\tilde{U}_{k+h,j})/|H_j| - \sigma_h\right| &\leq \Theta_4^3\left(16\Delta_4(t_n/3+1) + 6\Theta_4\sqrt{t_n/l_n} + 4\Psi_4(t_n/3+1)\right) \\ &\leq C\left(s_n^{-\alpha^2\iota/(1+\alpha)} + n^{-(1-\iota)\gamma/2}\right). \end{aligned}$$

Choose ℓ such that $0 < \ell < \min\{(1-\eta)(p-4)/8, \alpha\beta, \alpha^2\iota/(1+\alpha), (1-\iota)\gamma/2, \gamma-\beta\}$. Then

$$\begin{aligned} |\text{Cov}(\mathcal{R}_{n,k}, \mathcal{R}_{n,k+h})/n - \sigma_h| &\leq C_p\left(n^{-(1-\eta)(p-4)/8}(\log n)^{(p-4)/4} + n^{-\alpha\beta}\right. \\ &\quad \left.+ s_n^{-\alpha^2\iota/(1+\alpha)} + n^{-(1-\iota)\gamma/2}\right) + \frac{2w_n m_n \sigma_0}{n} \leq C_p s_n^{-\ell} \end{aligned}$$

and the lemma follows. \square

4.2.5. Step 5: Moderate deviations.

Proof of Lemma 14. Note that for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $|\mathbf{x} + \mathbf{y}|_\bullet \leq |\mathbf{x}|_\bullet + |\mathbf{y}|_\bullet$. Let $\mathbf{Z} \sim \mathcal{N}(0, I_d)$ and $\theta_n = (\log s_n)^{-1}$. Since $|\bar{U}_{k,j}| \leq 2\sqrt{n}/(\log s_n)^3$, by Fact 2.2 of Einmahl and Mason (1997),

$$\begin{aligned} P(|\mathcal{R}_n/\sqrt{n}|_\bullet \geq z_n) &\leq P(|\Sigma_n^{1/2}\mathbf{Z}|_\bullet \geq z_n - \theta_n) + P(|\mathcal{R}_n/\sqrt{n} - \Sigma_n^{1/2}\mathbf{Z}| \geq \theta_n) \\ &\leq P(|\Sigma_n^{1/2}\mathbf{Z}|_\bullet \geq z_n - \theta_n) + C_{p,d} \exp\left\{-C_{p,d}^{-1}(\log s_n)^2\right\}. \end{aligned}$$

By Lemma 23, the smallest eigenvalue of Σ is bounded from below by some $c_d > 0$ uniformly on $1 \leq k_1 < k_2 < \dots < k_d$. By Lemma 13 we have $\rho(\Sigma_n^{1/2} - \Sigma^{1/2}) \leq c_d^{-1/2} \cdot \rho(\Sigma_n - \Sigma) \leq C_{p,d} s_n^{-\ell}$, where the first inequality is taken from Problem 7.2.17 of Horn and Johnson (1990). It follows that by (74) and elementary calculations that

$$\begin{aligned} P(|\Sigma_n^{1/2} \mathbf{Z}|_{\bullet} \geq z_n - \theta_n) &\leq P(|\Sigma^{1/2} \mathbf{Z}|_{\bullet} \geq z_n - 2\theta_n) + P\left[\left|\left(\Sigma_n^{1/2} - \Sigma^{1/2}\right) \mathbf{Z}\right| \geq \theta_n\right] \\ &\leq P(|\Sigma^{1/2} \mathbf{Z}|_{\bullet} \geq z_n - 2\theta_n) + C_{p,d} \exp\{s_n^{-\ell}\}. \end{aligned}$$

By Lemma 22, we have

$$P(|\Sigma^{1/2} \mathbf{Z}|_{\bullet} \geq z_n - 2\theta_n) \leq \left[1 + C_{p,d}(\log s_n)^{-1/2}\right] P(|\Sigma^{1/2} \mathbf{Z}|_{\bullet} \geq z_n).$$

Putting these pieces together and observing that \mathbf{V} and $\Sigma^{1/2} \mathbf{Z}$ have the same distribution, we have

$$P(|\mathcal{R}_n/\sqrt{n}|_{\bullet} \geq z_n) \leq \left[1 + C_{p,d}(\log s_n)^{-1/2}\right] P(|\mathbf{V}|_{\bullet} \geq z_n) + C_{p,d} \exp\left\{-C_{p,d}^{-1}(\log s_n)^2\right\},$$

which together with a similar lower bound completes the proof of Lemma 14. \square

4.2.6. Proof of Theorem 1

After these preparation steps, we are now ready to prove Theorem 1.

Proof of Theorem 1. Set $z_n = a_{2s_n} x + b_{2s_n}$. It suffices to show

$$\lim_{n \rightarrow \infty} P\left(\max_{t_n < k \leq s_n} |\mathcal{R}_k/\sqrt{n}| \leq \sqrt{\sigma_0} z_n\right) = \exp\{-\exp(-x)\}. \quad (58)$$

Without loss of generality assume $\sigma_0 = 1$. Define the events $A_k = \{G_k \geq z_n\}$ and $B_k = \{\mathcal{R}_k/\sqrt{n} \geq z_n\}$. Let

$$Q_{n,d} = \sum_{t_n < k_1 < \dots < k_d \leq s_n} P(A_{k_1} \cap \dots \cap A_{k_d}) \quad \text{and} \quad \tilde{Q}_{n,d} = \sum_{t_n < k_1 < \dots < k_d \leq s_n} P(B_{k_1} \cap \dots \cap B_{k_d}).$$

By the inclusion-exclusion formula, we know for any $q \geq 1$

$$\sum_{d=1}^{2q} (-1)^{d-1} \tilde{Q}_{n,d} \leq P\left(\max_{t_n < k \leq s_n} |\mathcal{R}_k/\sqrt{n}| \geq a_{2s_n} x + b_{2s_n}\right) \leq \sum_{d=1}^{2q-1} (-1)^{d-1} \tilde{Q}_{n,d}. \quad (59)$$

By Lemma 14, $|\tilde{Q}_{n,d} - Q_{n,d}| \leq C_{p,d}(\log s_n)^{-1/2} Q_{n,d} + s_n^{-1}$. By Lemma 20 with elementary calculations, we know $\lim_{n \rightarrow \infty} Q_{n,d} = e^{-dx}/d!$, and hence $\lim_{n \rightarrow \infty} \tilde{Q}_{n,d} = e^{-dx}/d!$. By letting n go to infinity first and then d go to infinity in (59), we obtain (58), and the proof is complete. \square

4.3. Proof of Theorem 2

Proof of Theorem 2. We start with an m -dependence approximation that is similar to the proof of Theorem 1.

Set $m_n = \lfloor n^\beta \rfloor$ for some $0 < \beta < 1$. Define $\tilde{X}_i = \mathcal{H}_{i-m_n} X_i$, $\tilde{\gamma}_k = \mathbb{E}(\tilde{X}_0 \tilde{X}_k)$, and $\tilde{R}_{n,k} = \sum_{i=k+1}^n (\tilde{X}_{i-k} \tilde{X}_i -$

$\tilde{\gamma}_k$). Similarly as the proof of Lemma 10, we have under the condition (14)

$$\max_{1 \leq k < n} |R_{n,k} - \tilde{R}_{n,k}| = o_P \left(\sqrt{n/\log n} \right).$$

For $\tilde{R}_{n,k}$, we consider two cases according to whether $k \geq 3m_n$ or not.

Case 1: $k \geq 3m_n$. We first split the interval $[k+1, n]$ into the following big blocks of size $(k - m_n)$

$$\begin{aligned} H_j &= [k + j - 1(k - m_n) + 1, k + j(k - m_n)] \quad \text{for } 1 \leq j \leq w_n - 1 \\ H_{w_n} &= [k + (w_n - 1)(k - m_n) + 1, n], \end{aligned}$$

where w_n is the smallest integer such that $k + w_n(k - m_n) \geq n$. For each block H_j , we further split it into small blocks of size $2m_n$

$$\begin{aligned} K_{j,l} &= [k + (j - 1)(k - m_n) + (l - 1)2m_n + 1, k + (j - 1)(k - m_n) + 2lm_n] \quad \text{for } 1 \leq l < v_j \\ K_{j,v_j} &= [k + (v_j - 1)(k - m_n) + (l - 1)2m_n + 1, k + (j - 1)(k - m_n) + |H_j|] \end{aligned}$$

where v_j is the smallest integer such that $2m_n v_j \geq |H_j|$. Now define $U_{k,j,l} = \sum_{i \in K_{j,l}} \tilde{X}_{i-k} \tilde{X}_i$ and

$$\tilde{R}_{n,k}^{u,1} = \sum_{j \equiv u \pmod{3}} \sum_{l \text{ odd}} U_{k,j,l} \quad \text{and} \quad \tilde{R}_{n,k}^{u,2} = \sum_{j \equiv u \pmod{3}} \sum_{l \text{ even}} U_{k,j,l} \quad (60)$$

for $u = 0, 1, 2$. Observe that each $\tilde{R}_{n,k}^{u,o}$ ($u = 0, 1, 2$; $o = 1, 2$) is a sum of independent random variables. By (35), $\|U_{k,j,l}\| \leq 2\kappa_4 \Theta_4 |U_{k,j,l}|^{1/2}$. By Corollary 1.7 of Nagaev (1979) where we take $y_i = \sqrt{n}$ in their result, we have for any $\lambda > 0$

$$\begin{aligned} P \left(|\tilde{R}_{n,k}| \geq 6\lambda \sqrt{n \log n} \right) &\leq \sum_{u=0}^2 \sum_{o=1,2} P \left(\left| \tilde{R}_{n,k}^{u,o} \right| \geq \lambda \sqrt{n \log n} \right) \\ &\leq \sum_{u=0}^2 \sum_{o=1,2} \sum_{j,l}^* P \left(|U_{k,j,l}| \geq \lambda \sqrt{n \log n} \right) + 12 \left(\frac{C_p n^{1-\beta} \cdot n^{\beta p/4}}{n^{p/4}} \right)^{p\sqrt{\log n}/(p+4)} \\ &\quad + 12 \exp \left\{ -\frac{2\lambda^2}{(p+4)^2 \cdot e^{p/2} \cdot \kappa_4^2 \cdot \Theta_4^2} \cdot \log n \right\} =: I_{n,k} + II_{n,k} + III_{n,k}, \end{aligned} \quad (61)$$

where the range of j, l in the sum $\sum_{j,l}^*$ is as in (60). Clearly, $\sum_{k=3m_n}^{n-1} II_{n,k} = o(1)$. Similarly as the proof of Lemma 12, we can show that $\sum_{k=3m_n}^{n-1} I_{n,k} = o(1)$. Therefore, if $\epsilon = c_p/6$, then $\sum_{k=3m_n}^{n-1} III_{n,k} = O(n^{-1})$.

Case 2: $1 \leq k < 3m_n$. This case is easier. By splitting the interval $[k+1, n]$ into blocks with size $4m_n$ and using a similar argument as (61), we have

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{3m_n-1} P \left(|\tilde{R}_{n,k}| \geq c_p \sqrt{n \log n} \right) = 0.$$

The proof is complete. □

4.4. Box-Pierce tests

Similarly as the proof of Theorem 1, we use m -dependence approximations and blocking arguments to prove Theorem 4. We first outline the intermediate steps and give the main proof in Section 4.4.1, and then provide proofs of the intermediate lemmas in Section 4.4.2 and Section 4.4.3. We prove Theorem 6 in Section 4.4.4, and prove Corollary 5 and 7 in Section 4.4.5.

4.4.1. Proof of Theorem 4

Step 1: m -dependence approximation. Recall that $R_{n,k} = \sum_{i=k+1}^n (X_{i-k}X_i - \gamma_k)$. Without loss of generality, assume $s_n \leq \lfloor n^\beta \rfloor$. Set $m_n = 2\lfloor n^\beta \rfloor$. Let $\tilde{X}_i = \mathcal{H}_{i-m_n}^i X_i$ and $\tilde{R}_{n,k} = \sum_{i=k+1}^n (\tilde{X}_{i-k}\tilde{X}_i - \tilde{\gamma}_k)$. By (35) and (43), we know if $\Theta_4(m) = o(m^{-\alpha})$ for some $\alpha > 0$, then for all $1 \leq k \leq s_n$

$$\mathbb{E}|R_{n,k}^2 - \tilde{R}_{n,k}^2| \leq \|R_{n,k} + \tilde{R}_{n,k}\| \cdot \|R_{n,k} - \tilde{R}_{n,k}\| \leq C \Theta_4^3 \cdot n \cdot \Theta_4(m_n/2) = o(n^{1-\alpha\beta}).$$

The condition $\sum_{k=0}^\infty k^6 \delta_8(k) < \infty$ implies that $\Theta_4(m) = O(m^{-6})$. Therefore, under the conditions of Theorem 4, we have

$$\frac{1}{n\sqrt{s_n}} \sum_{k=1}^{s_n} \mathbb{E}_0 \left(R_{n,k}^2 - \tilde{R}_{n,k}^2 \right) = o_P(1).$$

Step 2: Throw out small blocks. Let $l_n = \lfloor n^\eta \rfloor$, where $\eta \in (\beta, 1)$. Split the interval $[1, n]$ into alternating small and large blocks similarly as (41):

$$K_0 = [1, s_n]$$

$$H_j = [s_n + (j-1)(2m_n + l_n) + 1, s_n + (j-1)(2m_n + l_n) + l_n] \quad 1 \leq j \leq w_n$$

$$K_j = [s_n + (j-1)(2m_n + l_n) + l_n + 1, s_n + j(2m_n + l_n)]; \quad 1 \leq j \leq w_n - 1; \quad \text{and}$$

$$K_{w_n} = [s_n + (w_n - 1)(2m_n + l_n) + l_n + 1, n],$$

where w_n is the largest integer such that $s_n + (w_n - 1)(2m_n + l_n) + l_n \leq n$. Define $U_{k,0} = 0$, $V_{k,0} = \sum_{i \in K_0, i > k} (\tilde{X}_{i-k}\tilde{X}_i - \tilde{\gamma}_k)$, and $U_{k,j} = \sum_{i \in H_j} (\tilde{X}_{i-k}\tilde{X}_i - \tilde{\gamma}_k)$, $V_{k,j} = \sum_{i \in K_j} (\tilde{X}_{i-k}\tilde{X}_i - \tilde{\gamma}_k)$ for $1 \leq j \leq w_n$. Set $\mathcal{R}_{n,k} = \sum_{j=1}^{w_n} U_{k,j}$. Observe that by construction, $U_{k,j}$, $1 \leq j \leq w_n$ are iid random variables. In the following lemma we show that it suffices to consider $\mathcal{R}_{n,k}$.

Lemma 15. Assume $X_i \in \mathcal{L}^8$, $\mathbb{E}X_i = 0$, and $\sum_{k=0}^\infty k^6 \delta_8(k) < \infty$, then

$$\frac{1}{n\sqrt{s_n}} \sum_{k=1}^{s_n} \mathbb{E}_0 \left(\tilde{R}_{n,k}^2 - \mathcal{R}_{n,k}^2 \right) = o_P(1).$$

Step 3: Central limit theorem concerning $\mathcal{R}_{n,k}$'s.

Lemma 16. Assume $X_i \in \mathcal{L}^8$, $\mathbb{E}X_i = 0$, and $\sum_{k=0}^\infty k^6 \delta_8(k) < \infty$, then

$$\frac{1}{n\sqrt{s_n}} \sum_{k=1}^{s_n} (\mathcal{R}_{n,k}^2 - \mathbb{E}\mathcal{R}_{n,k}^2) \Rightarrow \mathcal{N} \left(0, 2 \sum_{k \in \mathbb{Z}} \sigma_k^2 \right).$$

We are now ready to prove Theorem 4.

Proof of Theorem 4. By Lemma 15 and Lemma 16, we know

$$\frac{1}{n\sqrt{s_n}} \sum_{k=1}^{s_n} (R_{n,k}^2 - \mathbb{E}R_{n,k}^2) \Rightarrow \mathcal{N}\left(0, 2 \sum_{k \in \mathbb{Z}} \sigma_k^2\right).$$

It remains to show that

$$\lim_{n \rightarrow \infty} \frac{1}{n\sqrt{s_n}} \sum_{k=1}^{s_n} [\mathbb{E}R_{n,k}^2 - (n-k)\sigma_0] = 0. \quad (62)$$

We need Lemma 24 with a slight modification. Observe that in equation (91), we now have $\sum_{j=1}^{m_n} \Theta_2(j)^2 < \infty$, and hence

$$|\mathbb{E}R_{n,k}^2 - (n-k)\sigma_0| \leq C \left[(n-k)\Delta_4(\lfloor k/3 \rfloor + 1) + \sqrt{n-k} \right]$$

With the condition $\Theta_8(m) = o(m^{-6})$, elementary calculations show that $\Delta_4(m) = o(m^{-5})$. Then (62) follows, and the proof is complete. \square

4.4.2. Step 2: Throw out small blocks.

Let \mathcal{A}_2 be the collection of all double arrays $A = (a_{ij})_{i,j \geq 1}$ such that

$$\|A\|_\infty := \max \left\{ \sup_{i \geq 1} \sum_{j=1}^{\infty} |a_{ij}|, \sup_{j \geq 1} \sum_{i=1}^{\infty} |a_{ij}| \right\} < \infty.$$

For $A, B \in \mathcal{A}_2$, define $AB = (\sum_{k=1}^{\infty} a_{ik}b_{kj})$. It is easily seen that $AB \in \mathcal{A}_2$ and $\|AB\|_\infty \leq \|A\|_\infty \|B\|_\infty$. Furthermore, this fact implies the following proposition, which will be useful in computing sums of products of cumulants. For $d \geq 0$, let \mathcal{A}_d be the collection of all d -dimensional array $A = A(i_1, i_2, \dots, i_d)$ such that

$$\|A\|_\infty := \max_{1 \leq j \leq d} \left\{ \sup_{i_j \geq 1} \sum_{\{i_k: k \neq j\}} |A(i_1, i_2, \dots, i_d)| \right\} < \infty.$$

Note that $\mathcal{A}_0 = \mathbb{R}$, and $\|A\|_\infty = |A|$ if $A \in \mathcal{A}_0$.

Proposition 17. For $k \geq 0$, $l \geq 0$ and $d \geq 1$, if $A \in \mathcal{A}_{k+d}$ and $B \in \mathcal{A}_{l+d}$, define an array C by

$$C(i_1, \dots, i_k, i_{k+1}, \dots, i_{k+l}) = \sum_{j_1, \dots, j_d \geq 1} A(i_1, \dots, i_k, j_1, \dots, j_d) B(j_1, \dots, j_d, i_{k+1}, \dots, i_{k+l})$$

then $C \in \mathcal{A}_{k+l}$, and $\|C\|_\infty \leq \|A\|_\infty \|B\|_\infty$.

In Lemma 18 we present an upper bound for $\text{Cov}(R_{n,k}, R_{n,h})$. We formulate the lemma in a more general way for later uses in the proofs of Lemma 15 and Lemma 16. For a k -dimensional random vector (Y_1, \dots, Y_k)

such that $\|Y_i\|_k < \infty$ for $1 \leq i \leq k$, denote by $\text{Cum}(Y_1, \dots, Y_k)$ its k -th order joint cumulant. For the stationary process $(X_i)_{i \in \mathbb{Z}}$, we write

$$\gamma(k_1, k_2, \dots, k_d) := \text{Cum}(X_0, X_{k_1}, X_{k_2}, \dots, X_{k_d}).$$

We need the assumption of summability of joint cumulants in Lemma 18, Lemma 15 and Lemma 16. For this reason, we provide a sufficient condition in Section 6.

Lemma 18. *Assume $X_i \in \mathcal{L}^4$, $\mathbb{E}X_i = 0$, $\Theta_2 < \infty$ and $\sum_{k_1, k_2, k_3 \in \mathbb{Z}} |\gamma(k_1, k_2, k_3)| < \infty$. For $k, h \geq 1$, $l_n \geq t_n > 0$ and $s_n \in \mathbb{Z}$, set $U_k = \sum_{i=1}^{l_n} (X_{i-k}X_i - \gamma_k)$ and $V_h = \sum_{j=s_n+1}^{s_n+t_n} (X_{j-h}X_j - \gamma_h)$, then we have*

$$|\mathbb{E}(U_k V_h)| \leq t_n \Xi(k, h)$$

where $[\Xi(k, h)_{k, h \geq 1}]$ is a symmetric double array of non-negative numbers such that $\Xi \in \mathcal{A}_2$, and

$$\|\Xi\|_\infty \leq 2\Theta_2^4 + \sum_{k_1, k_2, k_3 \in \mathbb{Z}} |\gamma(k_1, k_2, k_3)|.$$

Proof. Write

$$\begin{aligned} \mathbb{E}(U_k V_h) &= \sum_{i=1}^{l_n} \sum_{j=1}^{t_n} \mathbb{E}[(X_{i-k}X_i - \gamma_k)(X_{s_n+j-h}X_{s_n+j} - \gamma_h)] \\ &= \sum_{i=1}^{l_n} \sum_{j=1}^{t_n} [\gamma(-k, j + s_n - i - h, j + s_n - i) \\ &\quad + \gamma_{j+s_n-i+k-h} \gamma_{j+s_n-i} + \gamma_{j+s_n-i+k} \gamma_{j+s_n-i-h}]. \end{aligned}$$

For the sum of the second term, we have

$$\begin{aligned} \left| \sum_{i=1}^{l_n} \sum_{j=1}^{t_n} \gamma_{j+s_n-i+k-h} \gamma_{j+s_n-i} \right| &= \left| \sum_{d=1}^{t_n-1} (\gamma_{s_n+d+k-h} \gamma_{s_n+d})(t_n - d) \right. \\ &\quad \left. + t_n \sum_{d=t_n-l_n}^0 \gamma_{s_n+d+k-h} \gamma_{s_n+d} \right. \\ &\quad \left. + \sum_{d=1-l_n}^{t_n-l_n-1} (\gamma_{s_n+d+k-h} \gamma_{s_n+d})(l_n + d) \right| \\ &\leq t_n \sum_{d \in \mathbb{Z}} |\gamma_{s_n+d+k-h} \gamma_{s_n+d}| \\ &\leq t_n \sum_{d \in \mathbb{Z}} \zeta_{d+k-h} \zeta_d. \end{aligned}$$

Similarly, for the sum of the last term

$$\left| \sum_{i=1}^{l_n} \sum_{j=1}^{t_n} \gamma_{j+s_n-i+k} \gamma_{j+s_n-i-h} \right| \leq t_n \sum_{d \in \mathbb{Z}} \zeta_{d+k+h} \zeta_d.$$

Observe that $\sum_{h=1}^{\infty} \sum_{d \in \mathbb{Z}} \zeta_{d+k-h} \zeta_d \leq (\sum_{d \in \mathbb{Z}} \zeta_d)^2 \leq \Theta_2^4$ and similarly $\sum_{h=1}^{\infty} \sum_{d \in \mathbb{Z}} \zeta_{d+k+h} \zeta_d \leq \Theta_2^4$. For the sum of the first term, it holds that

$$\left| \sum_{i=1}^{l_n} \sum_{j=1}^{t_n} \gamma(-k, j + s_n - i - h, j + s_n - i) \right| \leq t_n \sum_{d \in \mathbb{Z}} |\gamma(-k, d - h, d)|.$$

Utilizing the summability of cumulants, the proof is complete. \square

In the proof of Lemma 15, we need the concept of *indecomposable partitions*. Consider the table

$$\begin{array}{ccc} (1, 1) & \dots & (1, J_1) \\ \vdots & & \vdots \\ (I, 1) & \dots & (I, J_I) \end{array}$$

Denote the j -th row of the table by ϑ_j . A partition $\nu = \{\nu_1, \dots, \nu_q\}$ of the table is said to be *indecomposable* if there are no sets $\nu_{i_1}, \dots, \nu_{i_k}$ ($k < q$) and rows $\vartheta_{j_1}, \dots, \vartheta_{j_l}$ ($l < I$) such that $\nu_{i_1} \cup \dots \cup \nu_{i_k} = \vartheta_{j_1} \cup \dots \cup \vartheta_{j_l}$.

Proof of Lemma 15. Write

$$\begin{aligned} \sum_{k=1}^{s_n} \mathbb{E}_0(\tilde{R}_{n,k}^2 - \mathcal{R}_{n,k}^2) &= 2 \sum_{k=1}^{s_n} \mathbb{E}_0 \left[\mathcal{R}_{n,k} (\tilde{R}_{n,k} - \mathcal{R}_{n,k}) \right] + \sum_{k=1}^{s_n} \mathbb{E}_0 (\tilde{R}_{n,k} - \mathcal{R}_{n,k})^2 \\ &=: 2I_n + II_n. \end{aligned}$$

Using Lemma 16, we know $II_n/(n\sqrt{s_n}) = o_P(1)$. We can express I_n as

$$I_n = \sum_{a=0}^1 \sum_{b=0}^1 I_{n,ab} = I_{n,00} + I_{n,01} + I_{n,10} + I_{n,11}. \quad (63)$$

where for $a, b = 0, 1$ (assume without loss of generality that w_n is even),

$$I_{n,ab} = \sum_{k=1}^{s_n} \mathbb{E}_0 \left(\sum_{j=0}^{w_n/2} U_{k,2j-a} \sum_{j=0}^{w_n/2} V_{k,2j-b} \right).$$

Consider the first term in (63), write

$$\begin{aligned} \mathbb{E}(I_{n,00}^2) &= \sum_{k,h=1}^{s_n} \mathbb{E} \left[\sum_{j=1}^{w_n/2} \mathbb{E}_0(U_{k,2j} V_{k,2j}) \cdot \mathbb{E}_0(U_{h,2j} V_{h,2j}) \right] \\ &\quad + \sum_{k,h=1}^{s_n} \sum_{j_1 \neq j_2} \mathbb{E}(U_{k,2j_1} U_{h,2j_1}) \mathbb{E}(V_{k,2j_2} V_{h,2j_2}) \\ &\quad + \sum_{k,h=1}^{s_n} \sum_{j_1 \neq j_2} \mathbb{E}(U_{k,2j_1} V_{h,2j_1}) \mathbb{E}(V_{k,2j_2} U_{h,2j_2}) \\ &:= A_n + B_n + C_n. \end{aligned}$$

By Lemma 18, it holds that

$$|B_n| \leq \sum_{k,h=1}^{s_n} \sum_{j_1, j_2=0}^{w_n/2} l_n |K_{2j_2}| \cdot \left[\tilde{\Xi}(k, h) \right]^2$$

$$\leq w_n l_n \cdot (w_n m_n + 2l_n) \sum_{k,h=1}^{s_n} \left[\tilde{\Xi}_n(k, h) \right]^2 = o(n^2 s_n),$$

where $\tilde{\Xi}_n(k, h)$ is the $\Xi(k, h)$ (defined in Lemma 18) for the sequence (\tilde{X}_i) . Similarly,

$$\begin{aligned} |C_n| &\leq \sum_{k,h=1}^{s_n} \sum_{j_1, j_2=1}^{w_n/2} |K_{2j_1}| \cdot |K_{2j_2}| \cdot \left[\tilde{\Xi}_n(k, h) \right]^2 \\ &\leq (w_n m_n + l_n)^2 \sum_{k,h=1}^{s_n} \left[\tilde{\Xi}_n(k, h) \right]^2 = o(n^2 s_n). \end{aligned}$$

To deal with A_n , we express it in terms of cumulants

$$\begin{aligned} A_n &= \sum_{k,h=1}^{s_n} \sum_{j=1}^{w_n/2} [\text{Cum}(U_{k,2j}, V_{k,2j}, U_{h,2j}, V_{h,2j}) \\ &\quad + \mathbb{E}(U_{k,2j} U_{h,2j}) \mathbb{E}(V_{k,2j} V_{h,2j}) \\ &\quad + \mathbb{E}(U_{k,2j} V_{h,2j}) \mathbb{E}(V_{k,2j} U_{h,2j})] \\ &=: D_n + E_n + F_n. \end{aligned}$$

Apparently $|E_n| = o(n^2 s_n)$ and $|F_n| = o(n^2 s_n)$. Using the multilinearity of cumulants, we have

$$\text{Cum}(U_{k,2j}, V_{k,2j}, U_{h,2j}, V_{h,2j}) = \sum_{i_1, i_2 \in H_{2j}} \sum_{j_1, j_2 \in K_{2j}} \text{Cum}(\tilde{X}_{i_1-k} \tilde{X}_{i_1}, \tilde{X}_{j_1-k} \tilde{X}_{j_1}, \tilde{X}_{i_2-h} \tilde{X}_{i_2}, \tilde{X}_{j_2-h} \tilde{X}_{j_2})$$

for $1 \leq k, h \leq s_n$. By Theorem II.2 of Rosenblatt (1985), we know

$$\text{Cum}(\tilde{X}_{i_1-k} \tilde{X}_{i_1}, \tilde{X}_{j_1-k} \tilde{X}_{j_1}, \tilde{X}_{i_2-h} \tilde{X}_{i_2}, \tilde{X}_{j_2-h} \tilde{X}_{j_2}) = \sum_{\nu} \prod_{q=1}^b \text{Cum}(\tilde{X}_i, i \in \nu_q) \quad (64)$$

where the sum is over all indecomposable partitions $\nu = \{\nu_1, \dots, \nu_q\}$ of the table

$$\begin{array}{cc} i_1 - k & i_1 \\ j_1 - k & j_1 \\ i_2 - h & i_2 \\ j_2 - h & j_2 \end{array}$$

By Theorem 21, the condition $\sum_{k=0}^{\infty} k^6 \delta_8(k) < \infty$ implies that all the joint cumulants up to order eight are absolutely summable. Therefore, using Proposition 17, we know

$$\sum_{k,h=1}^{s_n} |\text{Cum}(U_{k,2j}, V_{k,2j}, U_{h,2j}, V_{h,2j})| = O(|K_{2j}| s_n^2),$$

and it follows that $|D_n| = O((w_n m_n + l_n) s_n^2) = o(n^2 s_n)$. We have shown that $\mathbb{E}(I_{n,00}^2) = o(n^2 s_n)$, which, in conjunction with similar results for the other three terms in (63), implies that $\mathbb{E}(I_n^2) = o(n^2 s_n)$ and hence $I_n/(n\sqrt{s_n}) = o_P(1)$. The proof is now complete. \square

4.4.3. Step 3: Central limit theorem concerning $\mathcal{R}_{n,k}$'s.

Proof of Lemma 16. Let $\Upsilon_n(k, h) := \mathbb{E}(U_{k,1}U_{h,1})$ and $v_n(k, h) := \Upsilon_n(k, h)/l_n$. By Lemma 18 we know $|v_n(k, h)| \leq \tilde{\Xi}_n(k, h)$. Write

$$\begin{aligned} \sum_{k=1}^{s_n} \mathbb{E}_0 \mathcal{R}_{n,k}^2 &= \sum_{k=1}^{s_n} \left[\sum_{j=1}^{w_n} (U_{k,j}^2 - \Upsilon_n(k, k)) + 2 \sum_{j=1}^{w_n} \left(U_{k,j} \sum_{l=1}^{j-1} U_{k,l} \right) \right] \\ &= \sum_{j=1}^{w_n} \left[\sum_{k=1}^{s_n} (U_{k,j}^2 - \Upsilon_n(k, k)) \right] + 2 \sum_{j=1}^{w_n} \left(\sum_{k=1}^{s_n} U_{k,j} \sum_{l=1}^{j-1} U_{k,l} \right). \end{aligned}$$

Using similar a argument as the one for dealing with the term A_n in Lemma 15, we know

$$\sum_{j=1}^{w_n} \left\| \sum_{k=1}^{s_n} (U_{k,j}^2 - \Upsilon_n(k, k)) \right\|^2 = o(n^2 s_n),$$

and it follows that

$$\frac{1}{n\sqrt{s_n}} \sum_{j=1}^{w_n} \left[\sum_{k=1}^{s_n} (U_{k,j}^2 - \Upsilon_n(k, k)) \right] = o_P(1).$$

Therefore, it suffices to consider

$$\sum_{j=1}^{w_n} \left(\sum_{k=1}^{s_n} U_{k,j} \sum_{l=1}^{j-1} U_{k,l} \right) =: \sum_{j=1}^{w_n} D_{n,j}.$$

Let $\mathcal{G}_{n,j} = \langle D_{n,1}, \dots, D_{n,j} \rangle$. Observe that $(D_{n,j})$ is a martingale difference sequence with respect to $(\mathcal{G}_{n,j})$.

We shall apply the martingale central limit theorem. Write

$$\begin{aligned} \mathbb{E}(D_{n,j}^2 | \mathcal{G}_{n,j-1}) - \mathbb{E} D_{n,j}^2 &= \sum_{k,h=1}^{s_n} \Upsilon_n(k, h) \left(\sum_{l=1}^{j-1} U_{k,l} \sum_{l=1}^{j-1} U_{h,l} - (j-1) \Upsilon_n(k, h) \right) \\ &= \sum_{k,h=1}^{s_n} \Upsilon_n(k, h) \left(\sum_{l=1}^{j-1} U_{k,l} U_{h,l} - (j-1) \Upsilon_n(k, h) \right) \\ &\quad + \sum_{k,h=1}^{s_n} \Upsilon_n(k, h) \left(\sum_{l=1}^{j-1} U_{k,l} \sum_{q=1}^{l-1} U_{h,q} + \sum_{l=1}^{j-1} U_{h,l} \sum_{q=1}^{l-1} U_{k,q} \right) \\ &=: I_{n,j} + II_{n,j} \end{aligned}$$

For the first term, by Lemma 18, we have

$$\begin{aligned} \left\| \sum_{j=1}^{w_n} I_{n,j} \right\|^2 &= \left\| \sum_{j=1}^{w_n-1} (w_n - j) \sum_{k,h=1}^{s_n} \Upsilon_n(k, h) [U_{k,j} U_{h,j} - \Upsilon_n(k, h)] \right\|^2 \\ &= \sum_{j=1}^{w_n-1} (w_n - j)^2 \left[\sum_{k,h} |\Upsilon_n(k, h)| \|(U_{k,j} U_{h,j} - \Upsilon_n(k, h))\| \right]^2 \\ &\leq w_n^3 l_n^4 \left[\sum_{k,h} |v_n(k, h)| \cdot 4\Theta_8^2 \right]^2 = o(n^4 s_n^2). \end{aligned}$$

Using Lemma 18 and Proposition 17, we obtain

$$\begin{aligned}
\left\| \sum_{j=1}^{w_n} \Pi_{n,j} \right\|^2 &= \left\| \sum_{j=1}^{w_n-1} (w_n - j) \sum_{k,h} \Upsilon_n(k,h) \left(U_{k,j} \sum_{l=1}^{j-1} U_{h,l} + U_{h,j} \sum_{l=1}^{j-1} U_{k,l} \right) \right\|^2 \\
&= 2 \sum_{j=1}^{w_n-1} (w_n - j)^2 (j-1) \sum_{1 \leq k_1, h_1, k_2, h_2 \leq s_n} \Upsilon_n(k_1, h_1) \Upsilon_n(k_2, h_2) [\Upsilon_n(k_1, k_2) \Upsilon_n(h_1, h_2) + \Upsilon_n(k_1, h_2) \Upsilon_n(h_1, k_2)] \\
&\leq 4n^4 \sum_{1 \leq k_1, h_1, k_2, h_2 \leq s_n} |v_n(k_1, h_1) v_n(h_1, h_2) v_n(h_2, k_2) v_n(k_2, k_1)| = O(n^4 s_n) = o(n^4 s_n^2).
\end{aligned}$$

Therefore, we have

$$\frac{1}{n^2 s_n} \left[\sum_{j=1}^{w_n} \mathbb{E} (D_{n,j}^2 | \mathcal{G}_{n,j-1}) - \sum_{j=1}^{w_n} \mathbb{E} D_{n,j}^2 \right] \xrightarrow{p} 0.$$

Using Lemma 18 and Lemma 24, we know

$$\frac{1}{n^2 s_n} \sum_{j=1}^{w_n} \mathbb{E} D_{n,j}^2 = \frac{1}{2n^2 s_n} w_n (w_n - 1) l_n^2 \sum_{k,h=1}^{s_n} [v_n(k, h)]^2 \rightarrow \frac{1}{2} \sum_{k \in \mathbb{Z}} \sigma_k^2,$$

and it follows that

$$\frac{1}{n^2 s_n} \sum_{j=1}^{w_n} \mathbb{E} (D_{n,j}^2 | \mathcal{G}_{n,j-1}) \xrightarrow{p} \frac{1}{2} \sum_{k \in \mathbb{Z}} \sigma_k^2. \tag{65}$$

To verify the Lindeberg condition, we compute

$$\begin{aligned}
\mathbb{E} D_{n,j}^4 &= \sum_{k_1, k_2, k_3, k_4=1}^{s_n} \mathbb{E} (U_{k_1,j} U_{k_2,j} U_{k_3,j} U_{k_4,j}) \\
&\quad \times \mathbb{E} \left[\left(\sum_{l=1}^{j-1} U_{k_1,l} \right) \left(\sum_{l=1}^{j-1} U_{k_2,l} \right) \left(\sum_{l=1}^{j-1} U_{k_3,l} \right) \left(\sum_{l=1}^{j-1} U_{k_4,l} \right) \right] \\
&\leq \sum_{k_1, k_2, k_3, k_4=1}^{s_n} |\mathbb{E} (U_{k_1,j} U_{k_2,j} U_{k_3,j} U_{k_4,j})| \cdot 2\mathcal{C}_4^4 (j-1)^2 l_n^2 \Theta_8^8
\end{aligned}$$

We express $\mathbb{E}(U_{k_1,1} U_{k_2,1} U_{k_3,1} U_{k_4,1})$ in terms of cumulants

$$\begin{aligned}
\mathbb{E}(U_{k_1,1} U_{k_2,1} U_{k_3,1} U_{k_4,1}) &= \text{Cum}(U_{k_1,1}, U_{k_2,1}, U_{k_3,1}, U_{k_4,1}) + \mathbb{E}(U_{k_1,1} U_{k_2,1}) \mathbb{E}(U_{k_3,1} U_{k_4,1}) \\
&\quad + \mathbb{E}(U_{k_1,1} U_{k_3,1}) \mathbb{E}(U_{k_2,1} U_{k_4,1}) + \mathbb{E}(U_{k_1,1} U_{k_4,1}) \mathbb{E}(U_{k_2,1} U_{k_3,1}) \\
&=: A_n + B_n + E_n + F_n
\end{aligned}$$

From Lemma 18, it is easily seen that

$$\sum_{k_1, k_2, k_3, k_4=1}^{s_n} |B_n| \leq l_n^2 \sum_{k_1, k_2, k_3, k_4=1}^{s_n} \tilde{\Xi}_n(k_1, k_2) \cdot \tilde{\Xi}_n(k_3, k_4) = O(l_n^2 s_n^2),$$

and similarly $\sum_{k_1, k_2, k_3, k_4=1}^{s_n} |E_n| = O(l_n^2 s_n^2)$ and $\sum_{k_1, k_2, k_3, k_4=1}^{s_n} |F_n| = O(l_n^2 s_n^2)$. By multilinearity of cumulants,

$$A_n = \sum_{i_1, i_2, i_3, i_4=1}^{l_n} \text{Cum}(\tilde{X}_{i_1-k_1} \tilde{X}_{i_1}, \tilde{X}_{i_2-k_2} \tilde{X}_{i_2}, \tilde{X}_{i_3-k_3} \tilde{X}_{i_3}, \tilde{X}_{i_4-k_4} \tilde{X}_{i_4}).$$

Each cumulant in the preceding equation is to be further simplified similarly as (64). Using summability of joint cumulants up to order eight and Proposition 17, we have

$$\sum_{k_1, k_2, k_3, k_4=1}^{s_n} |A_n| = O(l_n s_n^3) = o(l_n^2 s_n^2).$$

Using orders obtained for $|A_n|$, $|B_n|$, $|E_n|$ and $|F_n|$, we obtain $\sum_{j=1}^{w_n} \mathbb{E} D_{n,j}^4 = o(n^4 s_n^2)$. Then, by (65), we can apply Corollary 3.1. of Hall and Heyde (1980) to obtain

$$\frac{1}{n\sqrt{s_n}} \sum_{j=1}^{w_n} D_{n,j} \Rightarrow \mathcal{N}\left(0, \frac{1}{2} \sum_{k \in \mathbb{Z}} \sigma_k^2\right),$$

and the lemma follows. \square

4.4.4. Proof of Theorem 6

Proof of Theorem 6. We shall only prove (22), since (21) can be obtained by very similar arguments. Write $\hat{\gamma}_k = \mathbb{E}_0 \hat{\gamma}_k + \gamma_k - (\gamma_k - \mathbb{E} \hat{\gamma}_k)$, and hence

$$\begin{aligned} \sum_{k=1}^{s_n} (\hat{\gamma}_k^2 - \gamma_k^2) &= 2 \sum_{k=1}^{s_n} \gamma_k \mathbb{E}_0 \hat{\gamma}_k + \sum_{k=1}^{s_n} (\mathbb{E}_0 \hat{\gamma}_k)^2 - 2 \sum_{k=1}^{s_n} \frac{k}{n} \gamma_k \mathbb{E}_0 \hat{\gamma}_k - 2 \sum_{k=1}^{s_n} \frac{k}{n} \gamma_k^2 + \sum_{k=1}^{s_n} \frac{k^2}{n^2} \gamma_k^2 \\ &=: 2I_n + II_n + III_n + IV_n + V_n. \end{aligned}$$

Using the conditions $\Theta_4 < \infty$ and $s_n = o(\sqrt{n})$, it is easily seen that $\sqrt{n}IV_n \rightarrow 0$ and $\sqrt{n}V_n \rightarrow 0$. Furthermore

$$\sqrt{n}\|III_n\| \leq 2\sqrt{n} \sum_{k=1}^{s_n} \frac{k}{n} |\gamma_k| \cdot \frac{2\Theta_4^2}{\sqrt{n}} \rightarrow 0 \quad \text{and} \quad \sqrt{n}\mathbb{E}II_n \leq \sqrt{n} \sum_{k=1}^{s_n} \frac{4\Theta_4^4}{n} \rightarrow 0.$$

Define $Y_i = \sum_{k=1}^{\infty} \gamma_k X_{i-k}$. For the term I_n , write

$$\begin{aligned} nI_n &= \sum_{i=1}^n \mathbb{E}_0(X_i Y_i) - \sum_{i=1}^n \mathbb{E}_0 \left(X_i \sum_{k=s_n+1}^{\infty} \gamma_k X_{i-k} \right) + \sum_{k=1}^{s_n} \gamma_k \left(\sum_{i=1}^k (X_{i-k} X_i - \gamma_k) \right) \\ &=: A_n + B_n + E_n \end{aligned}$$

Clearly $\|E_n\|/\sqrt{n} \leq \sum_{k=1}^{s_n} |\gamma_k| 2\Theta_4^2 \sqrt{k}/\sqrt{n} \rightarrow 0$. Define $W_{n,i} = X_i \sum_{k=s_n+1}^{\infty} \gamma_k X_{i-k}$, then

$$\|\mathcal{P}^0 W_{n,i}\| \leq \begin{cases} \delta_4(i) \cdot \Theta_4 \sum_{k=s_n+1}^{\infty} |\gamma_k| & \text{if } 0 \leq i \leq s_n \\ \Theta_4 \delta_4(i) \sum_{k=s_n+1}^{\infty} |\gamma_k| + \Theta_4 \sum_{k=s_n+1}^i |\gamma_k| \delta_4(i-k) & \text{if } i > s_n. \end{cases}$$

It follows that

$$\|B_n/\sqrt{n}\| \leq 2\Theta_4^2 \sum_{k=s_n+1}^{\infty} |\gamma_k| \rightarrow 0.$$

Set $Z_i = X_i Y_i$, then (Z_i) is a stationary process of the form (9). Furthermore

$$\|\mathcal{P}^0 Z_i\| \leq \delta_4(i) \cdot \Theta_4 \sum_{k=1}^{\infty} |\gamma_k| + \Theta_4 \sum_{k=1}^i |\gamma_k| \delta_4(i-k).$$

Since $\sum_{i=0}^{\infty} \|\mathcal{P}^0 Z_i\| < \infty$, utilizing Theorem 1 in Hannan (1973) we have $A_n/\sqrt{n} \Rightarrow \mathcal{N}(0, \|D_0\|^2)$, and then (22) follows. \square

4.4.5. Proof of Corollary 5 and 7

Proof of Corollary 5 and 7. By (33), we know $\|n\bar{X}_n\|_4 \leq \sqrt{3n}\Theta_4$, and it follows that

$$\left\| \sum_{i=k+1}^n (X_{i-k} - \bar{X}_n)(X_i - \bar{X}_n) - \sum_{i=k+1}^n X_{i-k}X_i \right\| \leq 9\Theta_4^2.$$

Theorem 4 holds for $\check{\gamma}_k$ because

$$\begin{aligned} \frac{n}{\sqrt{s_n}} \sum_{k=1}^{s_n} \mathbb{E} |(\hat{\gamma}_k - \mathbb{E}\hat{\gamma}_k)^2 - (\check{\gamma}_k - \mathbb{E}\hat{\gamma}_k)^2| &\leq \frac{n}{\sqrt{s_n}} \sum_{k=1}^{s_n} \|\hat{\gamma}_k + \check{\gamma}_k - 2\mathbb{E}\hat{\gamma}_k\| \cdot \|\hat{\gamma}_k - \check{\gamma}_k\| \\ &\leq \frac{n}{\sqrt{s_n}} \sum_{k=1}^{s_n} \left(\frac{4\Theta_4^2}{\sqrt{n}} + \frac{9\Theta_4^2}{n} \right) \cdot \frac{9\Theta_4^2}{n} \rightarrow 0. \end{aligned}$$

In Theorem 6, (22) holds with $\hat{\gamma}_k$ replaced by $\check{\gamma}_k$ because

$$\sqrt{n} \sum_{k=1}^{s_n} \mathbb{E} |\hat{\gamma}_k^2 - \check{\gamma}_k^2| \leq \sqrt{n} \sum_{k=1}^{s_n} \|\hat{\gamma}_k + \check{\gamma}_k\| \cdot \|\hat{\gamma}_k - \check{\gamma}_k\| \leq \sqrt{n} \sum_{k=1}^{s_n} \left(2|\gamma_k| + \frac{4\Theta_4^2}{\sqrt{n}} + \frac{9\Theta_4^2}{n} \right) \frac{9\Theta_4^2}{n} \rightarrow 0,$$

and (21) can be proved similarly. Now we turn to the sample autocorrelations. Write

$$\sum_{k=1}^{s_n} \{[\hat{r}_k - (1 - k/n)r_k]^2 - [\hat{\gamma}_k/\gamma_0 - (1 - k/n)r_k]^2\} = \sum_{k=1}^{s_n} \frac{2(\mathbb{E}_0\hat{\gamma}_k)[\hat{\gamma}_k(\gamma_0 - \hat{\gamma}_0)]}{\gamma_0^2\hat{\gamma}_0} + \frac{\hat{\gamma}_k^2(\gamma_0 - \hat{\gamma}_0)^2}{\gamma_0^2\hat{\gamma}_0^2}.$$

Since

$$\sum_{k=1}^{s_n} \mathbb{E} |(\mathbb{E}_0\hat{\gamma}_k)\hat{\gamma}_k(\gamma_0 - \hat{\gamma}_0)| \leq \sum_{k=1}^{s_n} 2\mathcal{C}_3\Theta_6^2 \frac{1}{\sqrt{n}} \cdot \left(|\gamma_k| + 2\mathcal{C}_3\Theta_6^2 \frac{1}{\sqrt{n}} \right) \cdot 2\mathcal{C}_3\Theta_6^2 \frac{1}{\sqrt{n}} = o\left(\frac{\sqrt{s_n}}{n}\right)$$

and similarly $\sum_{k=1}^{s_n} \mathbb{E} |\hat{\gamma}_k^2(\gamma_0 - \hat{\gamma}_0)^2| = o(\sqrt{s_n}/n)$, (19) follows by applying the Slutsky theorem. To show the limit theorems in Corollary 7, note that using the Cramer-Wold device, we have

$$\left[\sqrt{n}(\hat{\gamma}_0^2 - \gamma_0^2), \sqrt{n} \left(\sum_{k=1}^{s_n} \hat{\gamma}_k^2 - \sum_{k=1}^{s_n} \gamma_k^2 \right) \right]$$

converges to a bivariate normal distribution. Then Corollary 7 follows by applying the delta method. \square

5. A Normal Comparison Principle

In this section we shall control tail probabilities of Gaussian vectors by using their covariance matrices. Denote by $\varphi_d((r_{ij}); x_1, \dots, x_d)$ the density of a d -dimensional multivariate normal random vector $\mathbf{X} = (X_1, \dots, X_d)^\top$ with mean zero and covariance matrix (r_{ij}) , where we always assume $r_{ii} = 1$ for $1 \leq i \leq d$ and (r_{ij}) is nonsingular. For $1 \leq h < l \leq d$, we use $\varphi_2((r_{ij}); X_h = x_h, X_l = x_l)$ to denote the marginal density of the sub-vector $(X_h, X_l)^\top$. Let

$$Q_d((r_{ij}); z_1, \dots, z_d) = \int_{z_1}^{\infty} \cdots \int_{z_d}^{\infty} \varphi_d((r_{ij}), x_1, \dots, x_d) \, dx_d \cdots dx_1.$$

The partial derivative with respect to r_{hl} is obtained similarly as equation (3.6) of Berman (1964) by using equation (3) of Plackett (1954)

$$\begin{aligned} & \frac{\partial Q_d((r_{ij}); z_1, \dots, z_d)}{\partial r_{hl}} \\ &= \left(\prod_{k \neq h, l} \int_{z_k}^{\infty} \right) \varphi_d((r_{ij}); x_1, \dots, x_{h-1}, z_h, x_{h+1}, \dots, x_{l-1}, z_l, x_{l+1}, \dots, x_d) \prod_{k \neq h, l} dx_k. \end{aligned} \quad (66)$$

where $\left(\prod_{k \neq h, l} \int_{z_k}^{\infty} \right)$ stands for $\int_{z_1}^{\infty} \cdots \int_{z_{h-1}}^{\infty} \int_{z_{h+1}}^{\infty} \cdots \int_{z_{l-1}}^{\infty} \int_{z_{l+1}}^{\infty} \cdots \int_{z_d}^{\infty}$. If all the z_k have the same value z , we use the simplified notation $Q_d((r_{ij}); z)$ and $\partial Q_d((r_{ij}); z)/\partial r_{hl}$. The following simple facts about conditional distribution will be useful. For four different indices $1 \leq h, l, k, m \leq d$, we have

$$\mathbb{E}(X_k | X_h = X_l = z) = \frac{r_{kh} + r_{kl}}{1 + r_{hl}} z, \quad (67)$$

$$\text{Var}(X_k | X_h = X_l = z) = \frac{1 - r_{hl}^2 - r_{kh}^2 - r_{kl}^2 + 2r_{hl}r_{kh}r_{kl}}{1 - r_{hl}^2}, \quad (68)$$

$$\text{Cov}(X_k, X_m | X_h = X_l = z) = r_{km} - \frac{r_{hk}r_{hm} + r_{lk}r_{lm} - r_{hl}r_{hk}r_{lm} - r_{hl}r_{hm}r_{lk}}{1 - r_{hl}^2}. \quad (69)$$

Lemma 19. *For every $z > 0$, $0 < s < 1$, $d \geq 1$ and $\epsilon > 0$, there exists positive constants C_d and ϵ_d such that for $0 < \epsilon < \epsilon_d$*

1. *if $|r_{ij}| < \epsilon$ for all $1 \leq i < j \leq d$, then*

$$Q_d((r_{ij}); z) \leq C_d \exp \left\{ - \left(\frac{d}{2} - C_d \epsilon \right) z^2 \right\} \quad (70)$$

$$Q_d((r_{ij}); z, \dots, z) \leq C_d f_d(\epsilon, 1/z) \exp \left\{ - \left(\frac{d}{2} - C_d \epsilon \right) z^2 \right\} \quad (71)$$

$$Q_d((r_{ij}); sz, z, \dots, z) \leq C_d \exp \left\{ - \left(\frac{s^2 + d - 1}{2} - C_d \epsilon \right) z^2 \right\} \quad (72)$$

where $f_{2k}(x, y) = \sum_{l=0}^k x^l y^{2(k-l)}$ and $f_{2k-1}(x, y) = \sum_{l=0}^{k-1} x^l y^{2(k-l)-1}$ for $k \geq 1$;

2. *if for all $1 \leq i < j \leq d+1$ such that $(i, j) \neq (1, 2)$, $|r_{ij}| \leq \epsilon$, then*

$$Q_{d+1}((r_{ij}); z) \leq C_d \exp \left\{ - \left(\frac{(1 - |r_{12}|)^2 + d}{2} - C_d \epsilon \right) z^2 \right\}. \quad (73)$$

Proof. The following facts about normal tail probabilities are well-known:

$$P(X_1 \geq x) \leq \frac{1}{\sqrt{2\pi}x} e^{-x^2/2} \text{ for } x > 0 \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{P(X_1 \geq x)}{(1/x)(2\pi)^{-1/2} \exp\{-x^2/2\}} = 1, \quad (74)$$

By (74), the inequalities (70) – (72) with $\epsilon = 0$ are true for the random vector with iid standard normal entries. The idea is to compare the desired probability with the corresponding one for such a vector. We first prove (70) by induction. When $d = 1$, the inequality is trivially true. When $d = 2$, by (66), there exists a number r'_{12} between 0 and r_{12} such that

$$|Q_2((r_{ij}); z) - Q_2(I_2; z)| \leq \varphi((r'_{ij}), z, z) |r_{12}|$$

$$\leq C \exp \left\{ -\frac{z^2}{1 + |r'_{12}|} \right\} \leq C \exp \{ -(1 - \epsilon)z^2 \},$$

which, together with $Q_2(I_2; z) \leq C \exp\{-z^2\}$, implies (70) for $d = 2$ with $\epsilon_2 = 1/2$ and some $C_2 > 1$. Now for $d \geq 3$, assume (70) holds for all dimensions less than d . There exists a matrix $(r'_{ij}) = \theta(r_{ij}) + (1 - \theta)I_d$ for some $0 < \theta < 1$ such that

$$Q_d((r_{ij}); z) - Q_d(I_d; z) = \sum_{1 \leq h, l \leq d} \frac{\partial Q_d}{\partial r_{hl}}((r'_{ij}); z, \dots, z) r_{hl}. \quad (75)$$

By (67), $\mathbb{E}(X_k | X_h = X_l = z) \leq 2\epsilon'z/(1 - \epsilon')$ for $k \neq h, l$. Therefore, by writing the density in (66) as the product of the density of (X_h, X_l) and the conditional density of $\mathbf{X}_{-\{h, l\}}$ given $X_h = X_l = z$, where $\mathbf{X}_{-\{h, l\}}$ denotes the sub-vector $(X_1, \dots, X_{h-1}, X_{h+1}, \dots, X_{l-1}, X_{l+1}, \dots, X_d)^\top$; we have

$$\left| \frac{\partial Q_d}{\partial r_{hl}}((r'_{ij}); z, \dots, z) \right| \leq \varphi_2((r'_{ij}); X_h = X_l = z) Q_{d-2}((r'_{ij|hl}); (1 - 3\epsilon)z), \quad (76)$$

where $(r'_{ij|hl})$ is the correlation matrix of the conditional distribution of $\mathbf{X}_{-\{h, l\}}$ given X_h and X_l . By (68) and (69), we know for $k, m \in [d] \setminus \{h, l\}$ and $k \neq m$,

$$\text{Var}(X_k | X_h = X_l = z) \geq 1 - 3\epsilon^2 - 2\epsilon^3 \quad \text{and} \quad \text{Cov}(X_k, X_m | X_h = X_l = z) \leq \frac{\epsilon(1 + \epsilon)}{1 - \epsilon}.$$

Therefore, all the off-diagonal entries of $(r'_{ij|hl})$ are less than 2ϵ if we let $\epsilon < 1/5$. Applying the induction hypothesis, if $2\epsilon < \epsilon_{d-2}$, then

$$Q_{d-2}((r'_{ij|hl}); (1 - 3\epsilon)z) \leq C_{d-2} \exp \left\{ -\left(\frac{d-2}{2} - 2C_{d-2}\epsilon \right) (1 - 3\epsilon)^2 z^2 \right\},$$

and equation (76) becomes

$$\left| \frac{\partial Q_d}{\partial r_{hl}}((r'_{ij}); z, \dots, z) \right| \leq C C_{d-2} \exp \{ -(1 - \epsilon)z^2 \} \cdot \exp \left\{ -\left(\frac{d-2}{2} - (2C_{d-2} + 3(d-2))\epsilon \right) z^2 \right\}.$$

Therefore, (70) holds for $\epsilon_d < \min\{1/5, \epsilon_{d-2}/2\}$ and some $C_d > 2C_{d-2} + 3(d-2) + 1$.

Using very similar arguments, inequality (72) can be proved by applying (70); and inequality (73) can be obtained by employing both (70) and (72). To prove inequality (71), which is a refinement of (70), it suffices to observe that, by (74), (75) and (76)

$$\begin{aligned} Q_d((r_{ij}); z) &\leq Q_d(I_d; z) + \sum_{1 \leq h, l \leq d} C \epsilon \exp\{-(1 - \epsilon)z^2\} Q_{d-2}((r'_{ij|hl}); (1 - 3\epsilon)z) \\ &\leq C_d \frac{1}{z^d} \exp \left\{ \frac{dz^2}{2} \right\} + C_d \epsilon \exp\{-(1 - \epsilon)z^2\} \sum_{1 \leq h, l \leq d} Q_{d-2}((r'_{ij|hl}); (1 - 3\epsilon)z); \end{aligned}$$

and apply the induction argument. \square

Lemma 20. *Let (X_n) be a stationary mean zero Gaussian process. Let $r_k = \text{Cov}(X_0, X_k)$. Assume $r_0 = 1$, and $\lim_{n \rightarrow \infty} r_n(\log n) = 0$. Let $a_n = (2 \log n)^{-1/2}$, $b_n = (2 \log n)^{1/2} - (8 \log n)^{-1/2}(\log \log n + \log 4\pi)$, and $z_n = a_n z + b_n$ for $z \in \mathbb{R}$. Define the event $A_i = \{X_i \geq z_n\}$, and*

$$Q_{n,d} = \sum_{1 \leq i_1 < \dots < i_d \leq n} P(A_{i_1} \cap \dots \cap A_{i_d}).$$

Then $\lim_{n \rightarrow \infty} Q_{n,d} = e^{-dz}/d!$ for all $d \geq 1$.

Proof. Note that $z_n^2 = 2 \log n - \log \log n - \log(4\pi) + 2z + o(1)$. If (X_n) consists of iid random variables, by the equality in (74),

$$\begin{aligned} \lim_{n \rightarrow \infty} Q_{n,d} &= \lim_{n \rightarrow \infty} \binom{n}{d} Q_d(I_d, z_n) \\ &= \lim_{n \rightarrow \infty} \binom{n}{d} \frac{1}{(2\pi)^{d/2} z_n^d} \exp \left\{ -\frac{dz_n^2}{2} \right\} = \frac{e^{-dz}}{d!}. \end{aligned}$$

When the X_n 's are dependent, the result is still trivially true when $d = 1$. Now we deal with the $d \geq 2$ case. Let $\gamma_k = \sup_{j \geq k} |r_j|$, then $\gamma_1 < 1$ by stationarity, and $\lim_{n \rightarrow \infty} \gamma_n \log n = 0$. Consider an ordered subset $J = \{t, t+l_1, t+l_1+l_2, \dots, t+l_1+\dots+l_{d-1}\} \subset [n]$, where $l_1, \dots, l_{d-1} \geq 1$. We define an equivalence relation \sim on J by saying $k \sim j$ if there exists $k_1, \dots, k_p \in J$ such that $k = k_1 < k_2 < \dots < k_p = j$, and $k_h - k_{h-1} \leq L$ for $2 \leq h \leq p$. For any $L \geq 2$, denote by $s(J, L)$ the number of l_j which are less than or equal to L . To simplify the notation, we sometimes use s instead of $s(J, L)$. J is divided into $d-s$ equivalence classes $\mathcal{B}_1, \dots, \mathcal{B}_{d-s}$. Suppose $s \geq 1$, assume w.l.o.g. that $|\mathcal{B}_1| \geq 2$. Pick $k_0, k_1 \in \mathcal{B}_1$, and $k_p \in \mathcal{B}_p$ for $2 \leq p \leq d-s$, and set $K = \{k_0, k_1, k_2, \dots, k_{d-s}\}$. Define $Q_J = P(\cap_{k \in J} A_k)$ and Q_K similarly, then $Q_J \leq Q_K$. By (73) of Lemma 19, there exists a number $M > 1$ depending on d and the sequence (γ_k) , such that when $L > M$,

$$\begin{aligned} Q_K &\leq C_{d-s} \exp \left\{ -\left(\frac{(1-\gamma_1)^2 + d-s}{2} - C_{d-s} \gamma_L \right) z_n^2 \right\} \\ &\leq C_{d-s} \exp \left\{ -\left(\frac{d-s}{2} + \frac{(1-\gamma_1)^2}{3} \right) z_n^2 \right\}. \end{aligned}$$

Note that $z_n^2 = 2 \log n - \log \log n + O(1)$. Pick $L_n = \max\{\lfloor n^\alpha \rfloor, M\}$ for some $\alpha < 2(1-\gamma_1^2)/3d$. For any $1 \leq a \leq d-1$, since there are at most $L_n^a n^{d-a}$ ordered subset $J \subset [n]$ such that $s(J, L_n) = a$, we know the sum of Q_J over these J is dominated by

$$C_{d-a} \exp \left\{ \log n \left((d-a) + \frac{2(d-1)(1-\gamma_1)^2}{3d} - (d-a) - \frac{2(1-\gamma_1)^2}{3} \right) \right\}$$

when n is large enough, which converges to zero. Therefore, it suffices to consider all the ordered subsets J such that $l_j > L_n$ for all $1 \leq j \leq d-1$.

Let $J = \{t_1, \dots, t_d\} \subset [n]$ be an ordered subset such that $t_i - t_{i-1} > L_n$ for $2 \leq i \leq d$, and $\mathcal{J}(d, L_n)$ be the collection of all such subsets. Let (r_{ij}) be the d -dimensional covariance matrix of \mathbf{X}_J . There exists a matrix $R_J = \theta(r_{ij})_{i,j \in J} + (1-\theta)I_d$ for some $0 < \theta < 1$ such that

$$Q_J - Q_d(I_d, z_n) = \sum_{h,l \in J, h < l} \frac{\partial Q_d}{\partial r_{hl}} [R_J; z_n] r_{ij}.$$

Let R_H , $H = J \setminus \{h, l\}$, be the correlation matrix of the conditional distribution of \mathbf{X}_H given X_h and X_l .

By (71) of Lemma 19, for n large enough

$$\begin{aligned}
\frac{\partial Q_d}{\partial r_{hl}}[R_J; z_n] &\leq C \exp \left\{ -\frac{z_n^2}{1 + \gamma_{l-h}} \right\} \cdot Q_{d-2}(R_K; (1 - 3\gamma_{L_n})z_n) \\
&\leq C C_{d-2} f_{d-2}(\gamma_{L_n}, 1/z_n) \exp \left\{ -\frac{z_n^2}{1 + \gamma_{l-h}} \right\} \\
&\quad \times \exp \left\{ -\left(\frac{d-2}{2} - 2C_{d-2}\gamma_{L_n} \right) (1 - 3\gamma_{L_n})^2 z_n^2 \right\} \\
&\leq C_d f_{d-2}(\gamma_{L_n}, 1/z_n) \exp \left\{ -\left(\frac{d}{2} - (2C_{d-2} + 3(d-2))\gamma_{L_n} - \gamma_{h-l} \right) z_n^2 \right\} \\
&\leq C_d f_{d-2}(\gamma_{L_n}, 1/z_n) \exp \left\{ -\left(\frac{d}{2} - C_d \gamma_{L_n} - \gamma_{h-l} \right) z_n^2 \right\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\sum_{J \in \mathcal{J}(d, L_n)} |Q_J - Q_d(I_d; z_n)| \\
&\leq C_d f_{d-2}(\gamma_{L_n}, 1/z_n) \sum_{J \in \mathcal{J}(d, L_n)} \sum_{1 \leq i < j \leq d} \exp \left\{ -\left(\frac{d}{2} - C_d \gamma_{L_n} - \gamma_{t_j - t_i} \right) z_n^2 \right\} \gamma_{t_j - t_i} \\
&= C_d f_{d-2}(\gamma_{L_n}, 1/z_n) \sum_{1 \leq i < j \leq d} \sum_{J \in \mathcal{J}(d, L_n)} \exp \left\{ -\left(\frac{d}{2} - C_d \gamma_{L_n} - \gamma_{t_j - t_i} \right) z_n^2 \right\} \gamma_{t_j - t_i}. \tag{77}
\end{aligned}$$

For each fixed pair $1 \leq i < j \leq d$, the inner sum in (77) is bounded by

$$\begin{aligned}
&C_d f_{d-2}(\gamma_{L_n}, 1/z_n) \sum_{l=L_n+1}^{n-1} (n-l)^{d-1} \exp \left\{ -\left(\frac{d}{2} - C_d \gamma_{L_n} - \gamma_l \right) z_n^2 \right\} \gamma_l \\
&\leq C_d f_{d-2}(\gamma_{L_n}, 1/z_n) (\log n)^{d/2} n^{-d} \sum_{l=L_n+1}^{n-1} (n-l)^{d-1} \exp \{ (C_d \gamma_{L_n} + \gamma_l) 2 \log n \} \gamma_l \tag{78}
\end{aligned}$$

$$\leq C_d f_{d-2}(\gamma_{\lfloor n^\alpha \rfloor}, 1/z_n) \gamma_{\lfloor n^\alpha \rfloor} (\log n)^{d/2} \exp \{ 2(C_d + 1) \gamma_{\lfloor n^\alpha \rfloor} \log n \}. \tag{79}$$

Since $\lim_{n \rightarrow \infty} \gamma_n \log n = 0$, it also holds that $\lim_{n \rightarrow \infty} \gamma_{\lfloor n^\alpha \rfloor} \log n = 0$. Note that $\lim_{n \rightarrow \infty} (\log n)^{1/2}/z_n = 2^{-1/2}$, it follows that $\lim_{n \rightarrow \infty} f_{d-2}(\gamma_{\lfloor n^\alpha \rfloor}, 1/z_n) (\log n)^{d/2-1} = 2^{-d/2+1}$. Therefore, the term in (79) converges to zero, and the proof is complete. \square

Remark 3. This lemma provides another proof of Theorem 3.1 in Berman (1964), which gives the asymptotic distribution of the maximum term of a stationary Gaussian process. They also showed that the theorem is true if the condition $\lim_{n \rightarrow \infty} r_n \log n = 0$ is replaced by $\sum_{n=1}^{\infty} r_n^2 < \infty$. Under the later condition, if we replace $\gamma_{t_j - t_i}$ by $|r_{t_j - t_i}|$ in (77), γ_l by $|r_l|$ in (78), then the term in (78) converges to zero, and hence our result remains true.

Remark 4. In the proof, the upper bounds on Q_J and $|Q_J - Q(I_d; z_n)|$ are expressed through the absolute values of the correlations, so we can obtain the same bounds for probabilities of the form $P(\cap_{1 \leq i \leq d} \{(-1)^{a_i} X_{t_i} \geq z_n\})$ for any $(a_1, \dots, a_d) \in \{0, 1\}^d$. Therefore, our result can be used to show the asymptotic distribution of

the maximum absolute term of a stationary Gaussian process. Specifically, we have

$$\lim_{n \rightarrow \infty} P \left(\max_{1 \leq i \leq n} |X_i| \leq a_{2n} x + b_{2n} \right) = \exp\{-\exp(-x)\}.$$

Deo (1972) obtained this result under the condition $\lim_{n \rightarrow \infty} r_n (\log n)^{2+\alpha} = 0$ for some $\alpha > 0$, whereas we only need $\lim_{n \rightarrow \infty} r_n \log n = 0$.

6. Summability of Cumulants

For a k -dimensional random vector (Y_1, \dots, Y_k) such that $\|Y_i\|_k < \infty$ for $1 \leq i \leq k$, the k -th order joint cumulant is defined as

$$\text{Cum}(Y_1, \dots, Y_k) = \sum (-1)^{p-1} (p-1)! \prod_{j=1}^p \left(\mathbb{E} \prod_{i \in \nu_j} Y_i \right), \quad (80)$$

where the summation extends over all partitions $\{\nu_1, \dots, \nu_p\}$ of the set $\{1, 2, \dots, k\}$ into p non-empty blocks. For a stationary process $(X_i)_{i \in \mathbb{Z}}$, we abbreviate

$$\gamma(k_1, k_2, \dots, k_d) := \text{Cum}(X_0, X_{k_1}, X_{k_2}, \dots, X_{k_d}),$$

Summability conditions of cumulants are often assumed in the spectral analysis of time series, see for example Brillinger (2001) and Rosenblatt (1985). Recently, such conditions were used by Anderson and Zeitouni (2008) in studying the spectral properties of banded sample covariance matrices. While such conditions are true for some Gaussian processes, functions of Gaussian processes (Rosenblatt, 1985), and linear processes with iid innovations (Anderson, 1971), they are not easy to verify in general. Wu and Shao (2004) showed that the summability of joint cumulants of order d holds under the condition that $\delta_d(k) = O(\rho^k)$ for some $0 < \rho < 1$. We present in Theorem 21 a generalization of their result. To simplify the proof, we introduce the composition of an integer. A *composition* of a positive integer n is an ordered sequence of strictly positive integers $\{v_1, v_2, \dots, v_q\}$ such that $v_1 + \dots + v_q = n$. Two sequences that differ in the order of their terms define different compositions. There are in total 2^{n-1} different compositions of the integer n . For example, we are giving in the following all of the eight compositions of the integer 4.

$$\{1, 1, 1, 1\} \quad \{1, 1, 2\} \quad \{1, 2, 1\} \quad \{1, 3\} \quad \{2, 1, 1\} \quad \{2, 2\} \quad \{3, 1\} \quad \{4\}.$$

Theorem 21. Assume $d \geq 2$, $X_i \in \mathcal{L}^{d+1}$ and $\mathbb{E}X_i = 0$. If

$$\sum_{k=0}^{\infty} k^{d-1} \delta_{d+1}(k) < \infty, \quad (81)$$

then

$$\sum_{k_1, \dots, k_d \in \mathbb{Z}} |\gamma(k_1, k_2, \dots, k_d)| < \infty. \quad (82)$$

Proof. By symmetry of the cumulant in its arguments and stationarity of the process, it suffices to show

$$\sum_{0 \leq k_1 \leq k_2 \leq \dots \leq k_d} |\gamma(k_1, k_2, \dots, k_d)| < \infty.$$

Set $X(k, j) := \mathcal{H}_j X_k$, we claim

$$\begin{aligned} \gamma(k_1, k_2, \dots, k_d) = \sum \text{Cum} [& X_0, X(k_1, 1), \dots, X(k_{v_1-1}, 1), X_{k_{v_1}} - X(k_{v_1}, 1), \\ & X(k_{v_1+1}, k_{v_1} + 1), \dots, X(k_{v_2-1}, k_{v_1} + 1), X_{k_{v_2}} - X(k_{v_2}, k_{v_1} + 1), \\ & \dots, \\ & X(k_{v_q+1}, k_{v_q} + 1), \dots, X(k_{d-1}, k_{v_q} + 1), X_{k_d} - X(k_d, k_{v_q} + 1)] ; \end{aligned} \quad (83)$$

where the sum is taken over all the 2^{d-1} increasing sequences $\{v_0, v_1, \dots, v_q, v_{q+1}\}$ such that $v_0 = 0, v_{q+1} = d$ and $\{v_1, v_2 - v_1, \dots, v_q - v_{q-1}, d - v_q\}$ is a composition of the integer d . We first consider the last summand which corresponds to the sequence $\{v_0 = 0, v_1 = d\}$,

$$\text{Cum} [X_0, X(k_1, 1), \dots, X(k_{d-1}, 1), X_{k_d} - X(k_d, 1)]$$

Observe that X_0 and $(X(k_1, 1), \dots, X(k_{d-1}, 1))$ are independent. By definition, only partitions for which X_0 and $X_{k_d} - X(k_d, 1)$ are in the same block contribute to the sum in (80). Suppose $\{\nu_1, \dots, \nu_p\}$ is a partition of the set $\{k_1, k_2, \dots, k_{d-1}\}$, since

$$\begin{aligned} \left| \mathbb{E} \left[X_0 (X_{k_d} - X(k_d, 1)) \prod_{k \in \nu_1} X(k, 1) \right] \right| &= \left| \sum_{j=-\infty}^0 \mathbb{E} \left[\mathcal{P}_j X_0 \mathcal{P}_j X_{k_d} \prod_{k \in \nu_1} X(k, 1) \right] \right| \\ &\leq \sum_{j=-\infty}^0 \delta_{d+1}(-j) \delta_{d+1}(k_d - j) \kappa_{d+1}^{|\nu_1|}, \end{aligned}$$

it follows that

$$\left| \mathbb{E} \left[X_0 (X_{k_d} - X(k_d, 1)) \prod_{k \in \nu_1} X(k, 1) \right] \cdot \prod_{j=2}^p \left(\mathbb{E} \prod_{k \in \nu_j} X(k, 1) \right) \right| \leq \sum_{j=0}^{\infty} \delta_{d+1}(j) \delta_{d+1}(k_d + j) \kappa_{d+1}^{d-1}$$

and therefore

$$\begin{aligned} & \sum_{0 \leq k_1 \leq k_2 \leq \dots \leq k_d} |\text{Cum} [X_0, X(k_1, 1), \dots, X(k_{d-1}, 1), X_{k_d} - X(k_d, 1)]| \\ & \leq C_d \sum_{0 \leq k_1 \leq k_2 \leq \dots \leq k_d} \sum_{j=0}^{\infty} \delta_{d+1}(j) \delta_{d+1}(k_d + j) \leq C_d \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \binom{k+d-1}{d-1} \delta_{d+1}(j) \delta_{d+1}(k+j) < \infty, \end{aligned}$$

provided that $\sum_{k=0}^{\infty} k^{d-1} \delta_{d+1}(k) < \infty$.

The other terms in (83) are easier to deal with. For example, for the term corresponding to the sequence $\{v_0 = 0, v_1 = 1, v_2 = d\}$, we have

$$\begin{aligned} & |\text{Cum} [X_0, X_{k_1} - X(k_1, 1), X(k_2, k_1 + 1), \dots, X(k_{d-1}, k_1 + 1), X_{k_d} - X(k_d, k_1 + 1)]| \\ & \leq C_d \kappa_{d+1}^{d-1} \Psi_{d+1}(k_1) \Psi_{d+1}(k_d - k_1). \end{aligned}$$

Since $\sum_{k=0}^{\infty} k^{d-1} \delta_{d+1}(k) < \infty$ implies $\sum_{k=0}^{\infty} k^{d-2} \Psi_{d+1}(k) \leq \infty$, it follows that

$$\begin{aligned} & \sum_{0 \leq k_1 \leq k_2 \leq \dots \leq k_d} |\text{Cum}[X_0, X_{k_1} - X(k_1, 1), X(k_2, k_1 + 1), \dots, X(k_{d-1}, k_1 + 1), X_{k_d} - X(k_d, k_1 + 1)]| \\ & \leq C_d \kappa_{d+1}^{d-1} \sum_{k=0}^{\infty} \Psi_{d+1}(k) \sum_{k=0}^{\infty} \binom{k+d-2}{d-2} \Psi_{d+1}(k) \leq \infty. \end{aligned}$$

We have shown that every cumulant in (83) is absolutely summable over $0 \leq k_1 \leq \dots \leq k_d$, and it remains to show the claim (83). We shall derive the case $d = 3$, (83) for other values of d are obtained using the same idea. By multilinearity of cumulants, we have

$$\begin{aligned} \gamma(k_1, k_2, k_3) &= \text{Cum}(X_0, X_{k_1}, X_{k_2}, X_{k_3}) \\ &= \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X_{k_2}, X_{k_3}] \\ &\quad + \text{Cum}[X_0, X(k_1, 1), X_{k_2} - X(k_2, 1), X_{k_3}] \\ &\quad + \text{Cum}[X_0, X(k_1, 1), X(k_2, 1), X_{k_3} - X(k_3, 1)] \\ &\quad + \text{Cum}[X_0, X(k_1, 1), X(k_2, 1), X(k_3, 1)]. \end{aligned}$$

Since X_0 and $(X(k_1, 1), X(k_2, 1), X(k_3, 1))$ are independent, the last cumulant is 0. Apply the same trick for the first two cumulants, we have

$$\begin{aligned} & \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X_{k_2}, X_{k_3}] \\ &= \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X_{k_2} - X(k_2, k_1 + 1), X_{k_3}] \\ &\quad + \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X(k_2, k_1 + 1), X_{k_3} - X(k_3, k_1 + 1)] \\ &\quad + \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X(k_2, k_1 + 1), X(k_3, k_1 + 1)] \\ &= \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X_{k_2} - X(k_2, k_1 + 1), X_{k_3} - X(k_3, k_2 + 1)] \\ &\quad + \text{Cum}[X_0, X_{k_1} - X(k_1, 1), X(k_2, k_1 + 1), X_{k_3} - X(k_3, k_1 + 1)] \end{aligned}$$

and

$$\text{Cum}[X_0, X(k_1, 1), X_{k_2} - X(k_2, 1), X_{k_3}] = \text{Cum}[X_0, X(k_1, 1), X_{k_2} - X(k_2, 1), X_{k_3} - X(k_3, k_2 + 1)].$$

Then the proof is complete. \square

Remark 5. When $d = 1$, (81) reduces to the *short-range dependence* or *short-memory* condition $\Theta_2 = \sum_{k=0}^{\infty} \delta_2(k) < \infty$. If $\Theta_2 = \infty$, then the process (X_i) may be long-memory in that the covariances are not summable. When $d \geq 2$, we conjecture that (81) can be weakened to $\Theta_{d+1} < \infty$. It holds for linear processes. Let $X_k = \sum_{i=0}^{\infty} a_i \epsilon_{k-i}$. Assume $\epsilon_k \in \mathcal{L}^{d+1}$ and $\sum_{k=0}^{\infty} |a_k| < \infty$, then $\delta_{d+1}(k) = |a_k| \|\epsilon_0\|_{d+1}$. Let $\text{Cum}_{d+1}(\epsilon_0)$ be the $(d+1)$ -th cumulant of ϵ_0 . Set $k_0 = 0$, by multilinearity of cumulants, we have

$$\gamma(k_1, \dots, k_d) = \sum_{t_0, t_1, \dots, t_d \geq 0} \left(\prod_{j=0}^d a_{t_j} \right) \text{Cum}(\epsilon_{-t_0}, \epsilon_{k_1-t_1}, \dots, \epsilon_{k_d-t_d})$$

40

$$= \sum_{t=0}^{\infty} \prod_{j=0}^d a_{k_j+t} \text{Cum}_{d+1}(\epsilon_0).$$

Therefore, the condition $\Theta_{d+1} < \infty$ suffices for (82). For a class of functionals of Gaussian processes, Rosenblatt (1985) showed that (82) holds if $\sum_{k=0}^{\infty} |\gamma_k| < \infty$, which in turn is implied by $\Theta_{d+1} < \infty$ under our setting. It is unclear whether in general the weaker condition $\Theta_{d+1} < \infty$ implies (82).

7. Some Auxiliary Lemmas

Suppose that \mathbf{X} is a d -dimensional random vector, and $\mathbf{X} \sim \mathcal{N}(0, \Sigma)$. If $\Sigma = I_d$, then by (74), it is easily seen that the ratio of $P(z_n - c_n \leq \|\mathbf{X}\|_{\bullet} \leq z_n)$ over $P(\|\mathbf{X}\|_{\bullet} \geq z_n)$ tends to zero provided that $c_n \rightarrow 0$, $z_n \rightarrow \infty$ and $c_n z_n \rightarrow 0$. It is a similar situation when Σ is not an identity matrix, as shown in the following lemma, which will be used in the proof of Lemma 14.

Lemma 22. *Let $\mathbf{X} \sim \mathcal{N}(0, \Sigma)$ be a d -dimensional normal random vector. Assume Σ is nonsingular. Let λ_0^2 and λ_1^2 be the smallest and largest eigenvalue of Σ respectively. Then for $0 < c < \delta < 1/2$ such that $A := (2\pi\lambda_1^2)^{(d-1)/2} \lambda_0^2 c^2 \delta^{-2} + d\delta \exp\{(\sqrt{6d}\lambda_1 + \lambda_0)/\lambda_0^3\} < 1$, then for any $z \in [1, \delta/c]$,*

$$P(z - c \leq \|\mathbf{X}\|_{\bullet} \leq z) \leq (1 - A)^{-1} A P(\|\mathbf{X}\|_{\bullet} \geq z). \quad (84)$$

Proof. Let $C_d = (6d)^{1/2} \lambda_1 / \lambda_0$. Since λ_0^2 is the smallest eigenvalue of Σ ,

$$\begin{aligned} P(\|\mathbf{X}\|_{\bullet} \geq z - c) &\geq (2\pi \det(\Sigma))^{-d/2} \exp\left\{-\frac{d(z+1)^2}{2\lambda_0^2}\right\} \\ &\geq (2\pi\lambda_1^2)^{-d/2} \exp\left\{-\frac{4d\delta^2}{2\lambda_0^2 c^2}\right\}. \end{aligned}$$

Since $P(\|\mathbf{X}\|_{\infty} \geq C_d \delta / c) \leq d(2\pi\lambda_1^2)^{-1/2} \exp\{6d\delta^2/(2\lambda_0^2 c^2)\}$, we have

$$P(\|\mathbf{X}\|_{\infty} \geq C_d \delta / c) \leq (2\pi\lambda_1^2)^{(d-1)/2} \lambda_0^2 c^2 \delta^{-2} P(\|\mathbf{X}\|_{\bullet} \geq z - c). \quad (85)$$

For $0 \leq k \leq \lfloor 1/\delta \rfloor$, define the orthotopes $R_k = [z + (k-1)c, z + kc] \times [z - c, C_d \delta / c]^{d-1}$. For two points $\mathbf{x} = (x_1, \dots, x_d) \in R_0$, $\mathbf{x}_k = (x_1 + kc, x_2, \dots, x_d) \in R_k$, we have $\mathbf{x}_k^{\top} \Sigma^{-1} \mathbf{x}_k - \mathbf{x}^{\top} \Sigma^{-1} \mathbf{x} \leq (2\sqrt{d}C_d + 1)/\lambda_0^2$, and hence $P(\mathbf{X} \in R_k) \geq \exp\{-(\sqrt{d}C_d + 1)/\lambda_0^2\} P(\mathbf{X} \in R_0)$ for any $1 \leq k \leq \lfloor 1/\delta \rfloor$. Since the same inequality holds for every coordinate, we have

$$P(z - c \leq \|\mathbf{X}\|_{\bullet} \leq z, \|\mathbf{X}\|_{\infty} \leq C_d \delta / c) \leq d\delta \exp\{(\sqrt{d}C_d + 1)/\lambda_0^2\} P(\|\mathbf{X}\|_{\bullet} \geq z - c) \quad (86)$$

Combine (85) and (86), we know $P(z - c \leq \|\mathbf{X}\|_{\bullet} \leq z) \leq A \cdot P(\|\mathbf{X}\|_{\bullet} \geq z - c)$. So (84) follows. \square

The preceding lemma requires the eigenvalues of Σ to be bounded both from above and away from zero. In our application, Σ is taken as the covariance matrix of $(G_{k_1}, G_{k_2}, \dots, G_{k_d})^{\top}$, where (G_k) is defined in (6). Furthermore, we need such bounds be uniform over all choices of $k_1 < k_2 < \dots < k_d$. Let $f(\omega) =$

$(2\pi)^{-1} \sum_{h \in \mathbb{Z}} \sigma_h \cos(h\omega)$ be the spectral density of (G_k) . A sufficient condition would be that there exists $0 < m < M$ such that

$$m \leq f(\omega) \leq M, \quad \text{for } \omega \in [0, 2\pi], \quad (87)$$

because the eigenvalues of the autocovariance matrix are bounded from above and below by the maximum and minimum values that f takes respectively. For the proof see Section 5.2 of Grenander and Szegő (1958). Clearly the upper bound in (87) is satisfied in our situation, because $\sum_{h \in \mathbb{Z}} |\sigma_h| < \infty$. However, the existence of lower bound in (87) rules out some classical times series models. For example, if (G_k) is the moving average of the form $G_k = (\eta_k + \eta_{k-1})/\sqrt{2}$, then $f(\omega) = (1 + \cos(\omega))/2\pi$, and $f(\pi) = 0$. Nevertheless, although the minimum eigenvalue of the autocovariance matrix converges to $\inf_{\omega \in [0, 2\pi]} f(\omega)$ as the dimension of the matrix goes to infinity, there does exist a positive lower bound for the smallest eigenvalues of all the principal submatrices with a fixed dimension.

Lemma 23. *If $\sum_{h \in \mathbb{Z}} \sigma_h^2 < \infty$, then for each $d \geq 1$, there exists a constant $C_d > 0$ such that*

$$\inf_{k_1 < k_2 < \dots < k_d} \lambda_{\min} \{ \text{Cov} [(G_{k_1}, G_{k_2}, \dots, G_{k_d})^\top] \} \geq C_d.$$

Proof. We use induction. It is clear that we can choose (C_d) to be a non-increasing sequence. Without loss of generality, let us assume $k_1 = 1$. The statement is trivially true when $d = 1$. Suppose it is true for all dimensions up to d , we now consider the dimension $(d + 1)$ case. There exist an integer N_d such that $\sum_{h=N_d} \sigma_h^2 < 2C_d^2/(d + 1)$. If all the differences $k_{i+1} - k_i \leq N_d$ for $1 \leq i \leq d - 1$, there are N_d^{d-1} possible choices of $k_1 = 1 < k_2 < \dots < k_d$. Since the process (G_k) is non-deterministic, for all these choices, the corresponding covariance matrices are non-singular. Pick $C'_d > 0$ to be the smallest eigenvalue of all these matrices. If there is one difference $k_{l+1} - k_l > N_d$, set $\Sigma_1 = \text{Cov}[(G_{k_i})_{1 \leq i \leq l}]$ and $\Sigma_2 = \text{Cov}[(G_{k_i})_{l < i \leq d}]$, then $\lambda_{\min}(\Sigma_1) \geq C_d$ and $\lambda_{\min}(\Sigma_2) \geq C_d$. It follows that for any real numbers c_1, c_2, \dots, c_d such that $\sum_{i=1}^d c_i^2 = 1$,

$$\begin{aligned} \sum_{1 \leq i, j \leq d} c_i c_j \text{Cov}(G_{k_i}, G_{k_j}) &= (c_1, \dots, c_i)^\top \Sigma_J(c_1, \dots, c_i) \\ &\quad + (c_{i+1}, \dots, c_d)^\top \Sigma_J(c_{i+1}, \dots, c_d) \\ &\quad + 2 \sum_{i \leq l, j > l} c_i c_j \sigma_{k_j - k_i} \\ &\geq C_d - 2 \left(\sum_{i \leq l, j > l} \sigma_{k_j - k_i}^2 \right)^{1/2} \left(\sum_{i \leq l, j > l} c_i^2 c_j^2 \right)^{1/2} \\ &\geq C_d - \frac{1}{2} \left(\frac{d+1}{2} \cdot \sum_{h=N_d} \sigma_h^2 \right)^{1/2} \geq \frac{C_d}{2}. \end{aligned}$$

Setting $C_{d+1} = \min\{C_d/2, C'_d\}$, the proof is complete. \square

The following lemma is used in the proof of Lemma 13.

Lemma 24. Assume $X_i \in \mathcal{L}^4$, $\mathbb{E}X_0 = 0$, and $\Theta_4 < \infty$. Assume $l_n \rightarrow \infty$, $k_n \rightarrow \infty$, $\check{m}_n < \lfloor k_n/3 \rfloor$ and $h \geq 0$. Define $S_{n,k} = \sum_{i=1}^{l_n} (X_{i-k}X_i - \gamma_k)$. Then

$$|\mathbb{E}(S_{n,k_n}S_{n,k_n+h})/l_n - \sigma_h| \leq \Theta_4^3 \left(16\Delta_4(\check{m}_n + 1) + 6\Theta_4\sqrt{\check{m}_n/l_n} + 4\Psi_4(\check{m}_n + 1) \right). \quad (88)$$

Proof. Let $\check{X}_i = \mathcal{H}_{i-\check{m}_n}^i X_i$, then \check{X}_i and \check{X}_{i-k_n} are independent, because $\check{m}_n \leq \lfloor k_n/3 \rfloor$. Define $\check{S}_{n,k} = \sum_{i=1}^{l_n} \check{X}_{i-k}\check{X}_i$. By (40), we have for any $k \geq 0$,

$$\left\| (S_{n,k} - \check{S}_{n,k})/\sqrt{l_n} \right\| \leq 4\kappa_4\Delta_4(\check{m}_n + 1). \quad (89)$$

By (35), $\|S_{n,k}/\sqrt{l_n}\| \leq 2\kappa_4\Theta_4$ for any $k \geq 0$, and it follows that

$$\begin{aligned} & |\mathbb{E}(S_{n,k_n}, S_{n,k_n+h}) - \mathbb{E}(\check{S}_{n,k_n}, \check{S}_{n,k_n+h})| \\ & \leq \|S_{n,k_n} - \check{S}_{n,k_n}\| \cdot \|S_{n,k_n+h}\| + \|\check{S}_{n,k_n}\| \cdot \|S_{n,k_n+h} - \check{S}_{n,k_n+h}\| \\ & \leq 16l_n\kappa_4^2\Theta_4\Delta_4(\check{m}_n + 1). \end{aligned} \quad (90)$$

For any $k > 3\check{m}_n$, define $M_{n,k} = \sum_{j=1}^{l_n} D_j$, where $D_j = \sum_{i=j}^{j+\check{m}_n} \check{X}_{i-k}\mathcal{P}^j\check{X}_i = \sum_{q=0}^{\check{m}_n} X_{j+q-k}\mathcal{P}^jX_{j+q}$. Observe that $\mathcal{P}^j\check{X}_{j+q}$ and \check{X}_{j+q-k} are independent, we have

$$\begin{aligned} \|\check{S}_{n,k} - M_{n,k}\| &= \left\| \sum_{i=1}^{l_n} \sum_{j=i-\check{m}_n}^i \check{X}_{i-k}\mathcal{P}^j\check{X}_i - \sum_{j=1}^{l_n} \sum_{i=j}^{j+\check{m}_n} \check{X}_{i-k}\mathcal{P}^j\check{X}_i \right\| \\ &\leq \left\| \sum_{j=1-\check{m}_n}^0 \sum_{i=1}^{j+\check{m}_n} \check{X}_{i-k}\mathcal{P}^j\check{X}_i \right\| + \left\| \sum_{j=l_n-\check{m}_n+1}^{l_n} \sum_{i=l_n+1}^{j+\check{m}_n} \check{X}_{i-k}\mathcal{P}^j\check{X}_i \right\| \\ &\leq 2 \left(\sum_{j=1}^{\check{m}_n} \kappa_2^2\Theta_2(j)^2 \right)^{1/2} \leq 2\kappa_2\Theta_2\sqrt{\check{m}_n} \end{aligned} \quad (91)$$

According to the proof of Theorem 2 of Wu (2009), when $k > 3\check{m}_n$ $\|M_{n,k}/\sqrt{n}\|^2 = \sum_{k \in \mathbb{Z}} \check{\gamma}_k^2$, where $\check{\gamma}_k = \mathbb{E}\check{X}_0\check{X}_k$. By (34) and (37), $|\check{\gamma}_k| \leq \zeta_k$; and hence

$$\begin{aligned} \|M_{n,k}/\sqrt{n}\|^2 &\leq \sum_{k \in \mathbb{Z}} \zeta_k^2 = \sum_{j,j'=0}^{\infty} \left(\delta_2(j)\delta_2(j') \sum_{k \in \mathbb{Z}} \delta_2(j+k)\delta_2(j'+k) \right) \\ &\leq \sum_{j,j'=0}^{\infty} \delta_2(j)\delta_2(j')\Psi_2^2 \leq \Theta_2^2\Psi_2^2. \end{aligned} \quad (92)$$

By (35) and (37), $\|\check{S}_{n,k}/\sqrt{l_n}\| \leq 2\kappa_4\Theta_4$ for any $k \geq 0$. Combining (91) and (92), we have

$$|\mathbb{E}(\check{S}_{n,k_n}, \check{S}_{n,k_n+h}) - \mathbb{E}(M_{n,k_n}, M_{n,k_n+h})| \leq (2\kappa_4\Theta_4 + \Theta_2\Psi_2)\sqrt{l_n} \cdot 2\kappa_2\Theta_2\sqrt{\check{m}_n}. \quad (93)$$

Observe that when $k_n > 3\check{m}_n$, $X_{q-k_n}X_{q'-k_n-h}$ and $\mathcal{P}^0X_q\mathcal{P}^0X_{q'}$ are independent for $0 \leq q, q' \leq \check{m}_n$.

Therefore,

$$\begin{aligned}
\mathbb{E}(M_{n,k_n} M_{n,k_n+h}) &= l_n \mathbb{E} \left(\sum_{q,q'=0}^{\tilde{m}_n} X_{q-k_n} X_{q'-k_n-h} \mathcal{P}^0 \check{X}_q \mathcal{P}^0 \check{X}_{q'} \right) \\
&= l_n \sum_{q,q'=0}^{\tilde{m}_n} \check{\gamma}_{q-q'+h} \mathbb{E} [(\mathcal{P}^0 \check{X}_q)(\mathcal{P}^0 \check{X}_{q'})] \\
&= l_n \sum_{k \in \mathbb{Z}} \check{\gamma}_{k+h} \sum_{q' \in \mathbb{Z}} \mathbb{E} [(\mathcal{P}^0 \check{X}_{q'+k})(\mathcal{P}^0 \check{X}_{q'})] \\
&= l_n \sum_{k \in \mathbb{Z}} \check{\gamma}_{k+h} \sum_{q' \in \mathbb{Z}} \mathbb{E} [(\mathcal{P}^{q'} \check{X}_k)(\mathcal{P}^{q'} \check{X}_0)] \\
&= l_n \sum_{k \in \mathbb{Z}} \check{\gamma}_{k+h} \check{\gamma}_k.
\end{aligned} \tag{94}$$

By (38), $|\gamma_k - \check{\gamma}_k| \leq 2\kappa_2 \Psi_2(m+1)$. Since $|\gamma_k| \leq \zeta_k$ and $|\check{\gamma}_k| \leq \zeta_k$, we have

$$\begin{aligned}
\left| \sigma_h - \sum_{k \in \mathbb{Z}} \check{\gamma}_{k+h} \check{\gamma}_k \right| &= \left| \sum_{k \in \mathbb{Z}} (\gamma_k \gamma_{k+h} - \check{\gamma}_k \check{\gamma}_{k+h}) \right| \\
&\leq 4\kappa_2 \Psi_2(m+1) \sum_{k \in \mathbb{Z}} \zeta_k \leq 4\kappa_2 \Psi_2(m+1) \Theta_2^2.
\end{aligned} \tag{95}$$

Combining (90), (93) and (95), the lemma follows by noting that κ_2, κ_4 are dominated by Θ_4 ; and $\Theta_2(\cdot)$, $\Psi_2(\cdot)$ and $\Psi_4(\cdot)$ are all dominated by $\Theta_4(\cdot)$. \square

References

- AN, H. Z., CHEN, Z. G. and HANNAN, E. J. (1982). Autocorrelation, autoregression and autoregressive approximation. *Ann. Statist.* **10** 926–936.
- ANDERSON, T. W. (1971). *The statistical analysis of time series*. John Wiley & Sons Inc., New York.
- ANDERSON, T. W. (1991). The asymptotic distributions of autoregressive coefficients Technical Report No. 26, Stanford University, Department of Statistics.
- ANDERSON, G. W. and ZEITOUNI, O. (2008). A CLT for regularized sample covariance matrices. *Ann. Statist.* **36** 2553–2576.
- ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1989). Two moments suffice for Poisson approximations: the Chen–Stein method. *Ann. Probab.* **17** 9–25.
- BERMAN, S. M. (1964). Limit theorems for the maximum term in stationary sequences. *Ann. Math. Statist.* **35** 502–516.
- BICKEL, P. J. and LEVINA, E. (2008a). Covariance regularization by thresholding. *Ann. Statist.* **36** 2577–2604.
- BICKEL, P. J. and LEVINA, E. (2008b). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227.
- BOX, G. E. P. and PIERCE, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.* **65** 1509–1526.
- BRILLINGER, D. R. (2001). *Time series. Classics in Applied Mathematics* **36**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. Data analysis and theory, Reprint of the 1981 edition.
- BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time series: theory and methods*, Second ed. *Springer Series in Statistics*. Springer-Verlag, New York.
- BURKHOLDER, D. L. (1988). Sharp inequalities for martingales and stochastic integrals. *Astérisque* 157-158 75–94. Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987).
- CAI, T. and JIANG, T. (2010). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices Technical Report, University of Pennsylvania and University of Minnesota.
- DAVISON, A. C. and HINKLEY, D. V. (1997). *Bootstrap methods and their application. Cambridge Series in Statistical and Probabilistic Mathematics* **1**. Cambridge University Press, Cambridge.

- DEO, C. M. (1972). Some limit theorems for maxima of absolute values of Gaussian sequences. *Sankhyā Ser. A* **34** 289–292.
- DEO, R. S. (2000). Spectral tests of the martingale hypothesis under conditional heteroscedasticity. *J. Econometrics* **99** 291–315.
- DUCHESNE, P., LI, L. and VANDERMEERSCHEN, J. (2010). On testing for serial correlation of unknown form using wavelet thresholding. *Computational Statistics and Data Analysis* **54** 2512–2531.
- DURBIN, J. and WATSON, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika* **37** 409–428.
- DURBIN, J. and WATSON, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika* **38** 159–178.
- DURLAUF, S. N. (1991). Spectral based testing of the martingale hypothesis. *J. Econometrics* **50** 355–376.
- EINMAHL, U. and MASON, D. M. (1997). Gaussian approximation of local empirical processes indexed by functions. *Probab. Theory Related Fields* **107** 283–311.
- ESCANCIANO, J. C. and LOBATO, I. N. (2009). An automatic Portmanteau test for serial correlation. *J. Econometrics* **151** 140–149.
- FAN, J. (1996). Test of significance based on wavelet thresholding and Neyman’s truncation. *J. Amer. Statist. Assoc.* **91** 674–688.
- GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz forms and their applications*. California Monographs in Mathematical Sciences. University of California Press, Berkeley.
- HAEUSLER, E. (1984). An exact rate of convergence in the functional central limit theorem for special martingale difference arrays. *Z. Wahrsch. Verw. Gebiete* **65** 523–534.
- HALL, P. (1979). On the rate of convergence of normal extremes. *J. Appl. Probab.* **16** 433–439.
- HALL, P. and HEYDE, C. C. (1980). *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics.
- HANNAN, E. J. (1973). Central limit theorems for time series regression. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **26** 157–170.
- HANNAN, E. J. (1974). The uniform convergence of autocovariances. *Ann. Statist.* **2** 803–806.
- HANNAN, E. J. and DEISTLER, M. (1988). *The statistical theory of linear systems*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.
- HANNAN, E. J. and HEYDE, C. C. (1972). On limit theorems for quadratic functions of discrete time series. *Ann. Math. Statist.* **43** 2058–2066.
- HONG, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* **64** 837–864.
- HONG, Y. and LEE, Y. J. (2003). Consistent testing for serial uncorrelation of unknown form under general conditional heteroscedasticity. Preprint, Cornell University, Department of Economics.
- HORN, R. A. and JOHNSON, C. R. (1990). *Matrix analysis*. Cambridge University Press, Cambridge. Corrected reprint of the 1985 original.
- HOROWITZ, J. L., LOBATO, I. N., NANKERVIS, J. C. and SAVIN, N. E. (2006). Bootstrapping the Box-Pierce Q test: A robust test of uncorrelatedness. *J. Econometrics* **133** 841–862.
- HOSOYA, Y. and TANIGUCHI, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.* **10** 132–153.
- JIANG, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14** 865–880.
- JIRAK, M. (2011). On the maximum of covariance estimators. *Journal of Multivariate Analysis* **102** 1032–1046.
- KÜNSCH, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241.
- LEE, J. and HONG, Y. (2001). Testing for serial correlation of unknown form using wavelet methods. *Econometric Theory* **17** 386–423.
- LIU, W.-D., LIN, Z. and SHAO, Q.-M. (2008). The asymptotic distribution and Berry-Esseen bound of a new test for independence in high dimension with an application to stochastic optimization. *Ann. Appl. Probab.* **18** 2337–2366.
- LIU, W. and WU, W. B. (2010). Asymptotics of spectral density estimates. *Econometric Theory* **26** 1218–1245.
- LJUNG, G. and BOX, G. E. P. (1978). Measure of lack of fit in time-series models. *Biometrika* **65** 297–303.
- NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Probab.* **7** 745–789.
- PHILLIPS, P. C. B. and SOLO, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20** 971–1001.
- PLACKETT, R. L. (1954). A reduction formula for normal multivariate integrals. *Biometrika* **41** 351–360.
- RIO, E. (2009). Moment inequalities for sums of dependent random variables under projective conditions. *J. Theoret. Probab.* **22** 146–163.
- ROBINSON, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple

- regression. *J. Econometrics* **47** 67–84.
- ROMANO, J. P. and THOMBS, L. A. (1996). Inference for autocorrelations under weak assumptions. *J. Amer. Statist. Assoc.* **91** 590–600.
- ROSENBLATT, M. (1985). *Stationary sequences and random fields*. Birkhäuser Boston Inc., Boston, MA.
- SCHOTT, J. R. (2005). Testing for complete independence in high dimensions. *Biometrika* **92** 951–956.
- SHAO, X. (2011). Testing for white noise under unknown dependence and its applications to diagnostic checking for time series models. *Econometric Theory* **FirstView** 1–32. <http://dx.doi.org/10.1017/S0266466610000253>.
- TONG, H. (1981). A note on a Markov bilinear stochastic process in discrete time. *J. Time Ser. Anal.* **2** 279–284.
- TONG, H. (1990). *Nonlinear time series. Oxford Statistical Science Series* **6**. The Clarendon Press Oxford University Press, New York. A dynamical system approach,.
- WIENER, N. (1958). *Nonlinear problems in random theory. Technology Press Research Monographs*. The Technology Press of The Massachusetts Institute of Technology and John Wiley & Sons, Inc., New York.
- WU, W. B. (2005). Nonlinear system theory: another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154 (electronic).
- WU, W. B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35** 2294–2320.
- WU, W. B. (2009). An asymptotic theory for sample covariances of Bernoulli shifts. *Stochastic Process. Appl.* **119** 453–467.
- WU, W. B. and SHAO, X. (2004). Limit theorems for iterated random functions. *J. Appl. Probab.* **41** 425–436.
- XIAO, H. and WU, W. B. (2010). Covariance matrix estimation for stationary time series. *preprint*.
- ZHOU, W. (2007). Asymptotic distribution of the largest off-diagonal entry of correlation matrices. *Trans. Amer. Math. Soc.* **359** 5345–5363.